

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:49:04 ; Search time 4552.99 Seconds
(without alignments)
9659.452 Million cell updates/sec

Title: US-10-759-889-1

Perfect score: 930

Sequence: 1 atggaacgactaaacgcat.....ggcggggaagaagaagtaa 930

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	930	100.0	110000	1	ECOW67_1	Continuation (2 of
2	930	100.0	110000	1	U0096_33	Continuation (34 o
3	915.6	98.5	307962	1	AP002584	AP002584 Escherich
4	915.2	98.4	11593	1	AE005552	AE005552 Escherich
5	914	98.3	292497	1	AE016989	AE016989 Shigella
6	912.4	98.1	11734	1	AE015339	AE015339 Shigella
7	882	94.8	300359	1	AE016767	AE016767 Escherich
8	675	72.6	212936	2	AC020970	AC020970 Mus muscu
9	670.8	72.1	258050	1	AL627278	AL627278 Salmonell
10	670.8	72.1	239991	1	AE016845	AE016845 Salmonell
11	666	71.6	21252	1	AE008855	AE008855 Salmonell
12	621.4	66.8	933	6	AR389774	AR389774 Sequence
13	550	59.1	110000	1	EX950851_03	Continuation (4 of
14	548	58.9	2788	1	ECOTLDD1	D4451 Escherichia
15	528.2	56.8	1512	1	AB077386	AB077386 Serratia
16	498.8	53.6	10553	1	AE013617	AE013617 Versinia
17	498.8	53.6	229338	1	AE017142	AE017142 Versinia
18	498.8	53.6	235050	1	AJ414158	AJ414158 Versinia
19	489.8	52.7	20157	1	AY647257	AY647257 Versinia

c	20	183.8	19.8	110000	2	AC020884_2	Continuation (3 of
c	21	170	18.3	290452	2	AC079167	AC079167 Mus muscu
	22	135	14.5	220050	1	AL646074	AL646074 Ralstonia
	23	131	14.1	348074	1	EX640449	EX640449 Bordetell
	24	131	14.1	349305	1	EX640433	EX640433 Bordetell
c	25	131	14.1	349726	1	EX640421	EX640421 Bordetell
	26	125.2	13.5	346547	1	AP003012	AP003012 Mesorhizo
c	27	124	13.3	305520	1	AE016780	AE016780 Pseudomon
c	28	123.6	13.3	11085	1	AE004450	AE004450 Pseudomon
c	29	123.4	13.3	10261	1	AE004445	AE004445 Pseudomon
	30	122	13.1	909	12	AY658944	AY658944 Synthetic
	31	121.8	13.1	921	12	AY658887	AY658887 Synthetic
c	32	119.6	12.9	348997	1	EX640427	EX640427 Bordetell
c	33	119.6	12.9	349008	1	EX640444	EX640444 Bordetell
c	34	117.6	12.6	300933	1	AE016791	AE016791 Pseudomon
c	35	117	12.6	11528	1	AE015885	AE015885 Shewanell
	36	116.8	12.6	346259	1	EX640435	EX640435 Bordetell
c	37	116.4	12.5	349028	1	EX640413	EX640413 Bordetell
c	38	115.2	12.4	346287	1	EX640450	EX640450 Bordetell
c	39	115.2	12.4	346359	1	EX640411	EX640411 Bordetell
c	40	115.2	12.4	348077	1	AP003000	AP003000 Mesorhizo
c	41	114.2	12.3	11773	1	AE012243	AE012243 Xanthomon
c	42	113.4	12.2	939	12	AY658980	AY658980 Synthetic
c	43	113.2	12.2	14748	1	AE011780	AE011780 Xanthomon
	44	113	12.2	10967	1	AE012487	AE012487 Xanthomon
	45	112.6	12.1	927	6	AR385349	AR385349 Sequence

ALIGNMENTS

RESULT 1
ECOW67_1
WPCOMMENT

Sequence split into 4 fragments LOCUS ECOW67 Accession U18997

Fragment Name	Begin	End
ECOW67_0	1	110000
ECOW67_1	100001	210000
ECOW67_2	200001	310000
ECOW67_3	300001	372438

Continuation (2 of 4) of ECOW67 from base 100001 (U18997 Escherichia coli K-12, chromosome

Query Match 100.0%; Score 930; DB 1; Length 110000;

Best Local Similarity 100.0%; Pred. No. 2.4e-230; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGAACGACTAAACGCACTCGGTGTTGGCCAAAGTAGTTGAATTTGGCTCTTTTACC	60
Db	70265	ATGGAACGACTAAACGCACTCGGTGTTGGCCAAAGTAGTTGAATTTGGCTCTTTTACC	70324
Qy	61	CCCGCCGCCACAGCTACAGATGAGGTTTCGTCCTCATCAGTCAGACGGTATCAAACTG	120
Db	70325	CCCGCCGCCACAGCTACAGATGAGGTTTCGTCCTCATCAGTCAGACGGTATCAAACTG	70384
Qy	121	GAAGATGATTGCGAGGTAAAGCTGTTAAACGCTAGCACACGACGATTTGGCTCACCGAA	180
Db	70385	GAAGATGATTGCGAGGTAAAGCTGTTAAACGCTAGCACACGACGATTTGGCTCACCGAA	70444
Qy	181	CCCGGTAGAAATTTACTACCGAGGTCGCGTCGTATGTTTCATGAAGTCGAGAGTGTTCAT	240
Db	70445	CCCGGTAGAAATTTACTACCGAGGTCGCGTCGTATGTTTCATGAAGTCGAGAGTGTTCAT	70504
Qy	241	GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTTCTTCA	300
Db	70505	GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTTCTTCA	70564
Qy	301	ACTATGGCACAAAATGTTCTCGCGGGGTGACACCCAAAATGCTGAAGAATATCCCAAGGT	360
Db	70565	ACTATGGCACAAAATGTTCTCGCGGGGTGACACCCAAAATGCTGAAGAATATCCCAAGGT	70624
Qy	361	TTGAGCGTCATCTGGTTACCGGAATTCACGCCCCCGACCTGATTCGCGAGGTCGTGGAT	420
Db	70625	TTGAGCGTCATCTGGTTACCGGAATTCACGCCCCCGACCTGATTCGCGAGGTCGTGGAT	70684

Qy 421 GTGGTATCCGCGTCGCGCGGTGACAGGATTCAGCCTCTTTTCCGCGCGTCTGGCGCG 480
Db 70685 GTGGTATCCGCGTCGCGCGGTGACAGGATTCAGCCTCTTTTCCGCGCGTCTGGCGCG 70744
Qy 481 ATCCCAATGCTGTGTCGCGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAA 540
Db 70745 ATCCCAATGCTGTGTCGCGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAA 70804
Qy 541 CCGCGCGATTTGAGTAGTCATTCATCGCTTGAAATACAGCGTCGCGCGCGCAATGAATTT 600
Db 70805 CCGCGCGATTTGAGTAGTCATTCATCGCTTGAAATACAGCGTCGCGCGCGCAATGAATTT 70864
Qy 601 GAATGATCCGACCGGAGGATCTGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
Db 70865 GAATGATCCGACCGGAGGATCTGACTCGCTGATCCCAAGGAAGATTTGTGACT 70924
Qy 661 AATGATCCGATGACGCTGTGCTGCTGACGCGGTCGCGGATCGCTACGTCGCG 720
Db 70925 AATGATCCGATGACGCTGTGCTGCTGACGCGGTCGCGGATCGCTACGTCGCG 70984
Qy 721 CTGATGCGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCCGTCGCGGTTAC 780
Db 70985 CTGATGCGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCCGTCGCGGTTAC 71044
Qy 781 CAGTCAGATCCAGCGCGGTTATGCGTTATATACCGGAAAGATTAAGCTGCGCTGAAG 840
Db 71045 CAGTCAGATCCAGCGCGGTTATGCGTTATATACCGGAAAGATTAAGCTGCGCTGAAG 71104
Qy 841 GTACAGTCGTGATCAACTCGCTGACGAGTATTTTGTGAGTCGCTAAATGTTTCAG 900
Db 71105 GTACAGTCGTGATCAACTCGCTGACGAGTATTTTGTGAGTCGCTAAATGTTTCAG 71164
Qy 901 GAGATCCAGCGCGCGGAAAGAGTAA 930
Db 71165 GAGATCCAGCGCGCGGAAAGAGTAA 71194

RESULT 2
U00096_33
WPCOMMENT
Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000

U00096_30 3100001 3110000
U00096_31 3200001 3210000
U00096_32 3300001 3310000
U00096_33 3400001 3410000
U00096_34 3500001 3510000
U00096_35 3600001 3610000
U00096_36 3700001 3710000
U00096_37 3800001 3810000
U00096_38 3900001 3910000
U00096_39 4000001 4010000
U00096_40 4100001 4110000
U00096_41 4200001 4210000
U00096_42 4300001 4310000
U00096_43 4400001 4410000
U00096_44 4500001 4510000
U00096_45 4600001 4610000
U00096_46 4600001 4639675
Continuation (34 of 47) of U00096 from base 3300001 (U00096 Escherichia coli K-12 MG1655
Query Match 100.0%; Score 930; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 2.4e-230;
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGAACGACTAAACGCAATGTCGGTGTGTCGCAAAAGTAGTTGAAATTTGGCTCTTTTACC 60
Db 87542 ATGGAACGACTAAACGCAATGTCGGTGTGTCGCAAAAGTAGTTGAAATTTGGCTCTTTTACC 87601
Qy 61 GCCGCCCCAGACAGCTACAGATGAGCGTTTCCTCATCATGACAGCGGTATCAAAACCTG 120
Db 87602 GCCGCCCCAGACAGCTACAGATGAGCGTTTCCTCATCATGACAGCGGTATCAAAACCTG 87661
Qy 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACACGACGATGCGCTGACCGAA 180
Db 87662 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACACGACGATGCGCTGACCGAA 87721
Qy 181 GCCGCTAGAAATTTACTACGAGGCTGCGTCTGATTCGTTTCATGAGTGCAGGATGTCAT 240
Db 87722 GCCGCTAGAAATTTACTACGAGGCTGCGTCTGATTCGTTTCATGAGTGCAGGATGTCAT 87781
Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGACCGCTACGCTTGGCTTTCTTCA 300
Db 87782 GAGCACTGTATGCTTCAATAACACCCCATCGGACCGCTACGCTTGGCTTTCTTCA 87841
Qy 301 ACTATGCCACAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360
Db 87842 ACTATGCCACAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 87901
Qy 361 TTGAGCGTCAATCTGGTTACCGGAATTTCCAGCCCGACCTGATTCGCGACGCTCTGGAT 420
Db 87902 TTGAGCGTCAATCTGGTTACCGGAATTTCCAGCCCGACCTGATTCGCGACGCTCTGGAT 87961
Qy 421 GTGCTGATCCGCTGCGCGGCTTTCAGAGATTTCAGGCTGTTTCGCGCGCTTGGCGCG 480
Db 87962 GTGCTGATCCGCTGCGCGGCTTTCAGAGATTTCAGGCTGTTTCGCGCGCTTGGCGCG 88021
Qy 481 ATGCCAATGCTGTGTCGCGCGGAAAGCTATCTACCAATACCGCATACCGGAAAAA 540
Db 88022 ATGCCAATGCTGTGTCGCGCGGAAAGCTATCTACCAATACCGCATACCGGAAAAA 88081
Qy 541 CCCGCCGATTTGAGTAGTCATTCATGGCTTGAATACAGCGTCGCGCGCGCAATGAATTT 600
Db 88082 CCCGCCGATTTGAGTAGTCATTCATGGCTTGAATACAGCGTCGCGCGCGCAATGAATTT 88141
Qy 601 GAATGATCCGACCGGAAAGGATCTGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
Db 88142 GAATGATCCGACCGGAAAGGATCTGACTCGCTGATCCCAAGGAAGATTTGTGACT 88201
Qy 661 AATGATCCGATGACGCTGTGTCGCTGATCCGCGGTCGCGGATCGCTACGTGCGCG 720
Db 88202 AATGATCCGATGACGCTGTGTCGCTGATCCGCGGTCGCGGATCGCTACGTGCGCG 88261
Qy 721 CTGATGCGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCTCTGTCGCGCGTTAC 780

Db 88262 CTGATCGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGTCGCCGTTTAC 88321
 QY 781 CAGTCAGATCCAGCCGCCGTTTATCGTTATATACGAAAGATAAGTCGCGTGAAG 840
 Db 88322 CAGTCAGATCCAGCCGCCGTTTATCGTTATATACGAAAGATAAGTCGCGTGAAG 88381
 QY 841 GTACAGGTGATCAACTCGCTGACGAGTATTTTGTGAGTGGTAAATTTGTTTCAAG 900
 Db 88382 GTACAGGTGATCAACTCGCTGACGAGTATTTTGTGAGTGGTAAATTTGTTTCAAG 88441
 QY 901 GAGATGCACGGCGCGGGAAGAGAGTAA 930
 Db 88442 GAGATGCACGGCGCGGGAAGAGAGTAA 88471

RESULT 3
 AP002564 307962 bp DNA linear BCT 20-MAR-2004
 LOCUS Escherichia coli O157:H7 DNA, complete genome, section 15/20.
 ACCESSION AP002564 BA000007
 VERSION AP002564.1 GI:13363382
 KEYWORDS Escherichia coli O157:H7
 SOURCE Escherichia coli O157:H7
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE
 1 Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S., Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T., Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and Shinagawa,H.
 Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak
 Genes Genet. Syst. 74 (5), 227-239 (1999)
 20198780
 10734605

TITLE
 Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak

JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED 20557356
 REFERENCE 11108008

AUTHORS
 3 Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S., Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T., Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and Shinagawa,H.
 Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak
 Gene 258 (1-2), 127-139 (2000)
 20564182
 11111050

TITLE
 Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12

JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED 21156231
 REFERENCE 11258796

AUTHORS
 5 (bases 1 to 307962)
 Hattori,M., Ishii,K. and Shiba,T.
 Direct Submission
 Submitted (26-JUN-2000) Masahira Hattori, Kitasato Institute for

COMMENT

Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan
 (E-mail:hattori@genome.ls.kitasato-u.ac.jp,
 URL:http://genome.ls.kitasato-u.ac.jp/, Tel:81-42-778-8194,
 Fax:81-42-778-8193)
 genome project
 This work was done in collaboration with Tetsuya Hayashi, Makoto Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Koza Makino, Ken Kurokawa, Katsushi Yokoyama, Masashi Tanaka, Takeshi Honda, Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata (Shinshu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe, Chihiro Sasakawa (University of Tokyo), Hideto Takami (Japan Marine Science and Technology Center), Naotake Ogasawara (Nara Institute of Science and Technology), Satoru Kuhara (Kyuushu University), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science.

FEATURES
 source

1..307962
 /organism="Escherichia coli O157:H7"
 /mol_type="genomic DNA"
 /strain="O157:H7"
 /sub_strain="RIMD 0509952"
 /db_xref="taxon:83334"
 72..653
 /gene="ECS3910"
 72..653
 /gene="ECS3910"
 /note="similar to MDAB_ECOLI_gil1789406 percent identity 100 in 193 aa (Conserved in E.coli K-12)"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="modulator of drug activity B"
 /protein_id="BAB37333.1"
 /db_xref="GI:13363383"
 /translation="MSNLIITINGAKFPAHNSGQINDLTTEVADGTLRLDGLHVDIVRA
 DSDYDKAEVQNFVADVIWQPMGMPVTKYIDDVFTGEGTLYASDGRTRK
 DPSKYGSGGLVGKKYMLSLTNAPMEAFTEKQDFHGVGDVGLVLPFHKAQFLGM
 EPLPTIANDVIRKPDVPRTEYRKHLVEIFG"
 684..998
 /gene="ECS3911"
 684..998
 /gene="ECS3911"
 /note="similar to YGIN_ECOLI_gil1789407 percent identity 100 in 104 aa (Conserved in E.coli K-12)"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB37334.1"
 /db_xref="GI:13363384"
 /translation="MLTVIAEIRPGQHHQAVLDQFAKIVPTVLKEGCHGVAPMV
 DCAAGVSFQSMAPDSIVMEQWESIAHLEAHLQTPHMKAYSEAVKGDVLEMNRILOP
 GI"

gene

CDS

gene

CDS

gene

CDS

complement (1043..1930)
 /gene="ECS3912"
 complement (1043..1930)
 /gene="ECS3912"
 /note="probable transcriptional regulator similar to regulators (Rpiir family) e.g. [Bacillus subtilis] gi|18248807|emb|CAB93068.1| percent identity 25 in 236 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="putative transcriptional regulator"
 /protein_id="BAB37335.1"
 /db_xref="GI:13363385"
 /translation="MIQOKTRVDIYGERFRFRASQLTFGLRAVASIYNEHREVVLDQ
 TAMEIATNTSDATVIRAIQALGFAGLRDLKLTLEQWLPALSSSEKSKSTTVSNLTS
 DVNTAIDFVLEGLHYTCNVLSEPNELAQAVALLVQARQVAIFGIGASGLADYTA
 RLFRNIGLPATALNRIGLAEOLIALQRGDVLIMQAQSAHREGTLTIREARRLGTP
 VIILTNALDSRFSKDSIVTHVPRGDEKGTPLHGTVLLCLEMIVMSVASVAFQRAVK
 TIKRINDFHRGLTGRKNG"

gene	complement(1927. .2874)	Best Local Similarity 99.0%; Pred. No. 1.4e-226;	
CDS	/gene="ECs3913"	Matches 921; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
	/complement(1927. .2874)		
	/gene="ECs3913"		
	/note="probable ferrichrome-binding protein, similar to ferrichrome-binding proteins e.g. [Bacillus subtilis] gi 1706795 sp P49936 FHUB_BACSU percent identity 27 in 220 aa"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="putative ferrichrome-binding protein"		
	/protein_id="BAB37336.1"		
	/db_xref="GI:13363386"		
	/translation="MRLFFSLILLSFFARATEPVQVTDLGRKVVTPAHPKRIVSLHDLDTIFLIEGVFPVASHGRPDGSHFIRSGALLTGVDVFNSSIAFTADIDIEAIVAAKPDLIITEPTNPPIERLEKIAFTVSDHLKGGAPBIYRLAELTGTQSLATLERRVQAOINALKLDQKTVSVIOANQKINVMHSYHSLGRVLRDAGFRFPPLIESTPEGRMDVSAERLPELDADVFVATWRGDTGKRPQDELAAMEKVMPCWCQFLTACRSGRVLSREASINSFASIGLMAQIQSIAGRPLPEAK"		
	/complement(2885. .3934)		
gene	/gene="ECs3914"		
CDS	/complement(2885. .3934)		
	/note="probable ferrichrome ABC transporter (permease) similar to ferrichrome ABC transporter (permease) e.g. [Bacillus subtilis] gi 1706795 sp P49937 FHUG_BACSU percent identity 33 in 319 aa"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="putative permease of ferrichrome ABC transporter"		
	/protein_id="BAB37337.1"		
	/db_xref="GI:13363387"		
	/translation="MSRIMNVGLRPLRVGKSTLVRPNVLVLLGGLFLFVAGLILFGLMHGFSFVASEVGRALFAPENVSTDIARYIVQDILRPVIMALLCGMLGWAGAAQSIARNGLADPGLIGKSGSVAVLWLIQFQPMLGFWRPVAGLAGLLNVALIVIFCARISRPFRVLIGVSWFPAAGIGVEMTTADVQVOTALMWLSGLHAANWLVGISACNM LPAALLLFTARTADIALGHQVATGLVNSRLALLRVAAPILITAVCVSCVGNIGFVGLRAPHRSRFLRGQGTLLGSAVSAGALLVADSIGRLAFPLQAPAGIISLIG GPFULLLWQRNSP"		
	/complement(3931. .4914)		
gene	/gene="ECs3915"		
CDS	/complement(3931. .4914)		
	/gene="ECs3915"		
	/note="probable ferrichrome ABC transporter (permease) similar to ferrichrome ABC transporters (permease) e.g. [Synecocystis sp.] gi 7442493 pir S74438 percent identity 43 in 315 aa, [Bacillus subtilis] gi 1706795 sp P49936 FHUB_BACSU percent identity 39"		
	/codon_start=1		
	/evidence=not experimental		
	/transl_table=11		
	/product="putative permease of ferrichrome ABC transporter"		
	/protein_id="BAB37338.1"		
	/db_xref="GI:13363388"		
	/translation="MWRGLTVTFIALLLVGVVPGIGARFIAPOTVLOAFLHFEPQNFHNIVIRLRLPRLAALLTGSLVAGALLQAVIRNPISGEPHILGNAGALAVYASALGLAFVGRPLASTGTGALLFLILLSSAGRSLTEPMKVTLCGVALSARVSSI DMAAILIDELQVLRWTLWGLADQWATIGTSAWFSLGGVLAITHLAPSLNMLALG DRMAQGLVSTLRTFTLLATALLCGAAVSIAGPIGVGLLVPOIVRRLVSDLRVL LPLSACVGAALLLADIIFARTLFTPELATGVTALVGAPVFMVIMATRMER"		
	/complement(4911. .5720)		
gene	/gene="ECs3916"		
CDS	/complement(4911. .5720)		
	/gene="ECs3916"		
	/note="probable ABC-type iron-siderophore transport system ATP-binding protein, similar to ABC-type iron-siderophore"		
Query Match	98.5%; Score 915.6; DB 1; Length 307962;		

RESULT 4
 AE005552
 LOCUS
 DEFINITION
 Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 171 of 290.


```

ACCESSION AB005552 AB005174
VERSION AB005552.1 GI:12517866
KEYWORDS
SOURCE
ORGANISM Escherichia coli O157:H7 EDL933
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 11593)
Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,
Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S.K.,
Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,
Welch, R.A., and Blattner, F.R.
Genome sequence of enterohemorrhagic Escherichia coli O157:H7
Nature 409 (6819), 529-533 (2001)
2 (bases 1 to 11593)
Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,
Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S.K.,
Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,
Welch, R.A., and Blattner, F.R.
Direct Submission
Submitted (22-OCT-2000) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. .11593
/organism="Escherichia coli O157:H7 EDL933"
/mol_type="genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/notes="enterohemorrhagic"
88. .1017
/genes="gseA"
/genes="gseA"
/genes="gseA"
/genes="gseA"
/genes="gseA"
/function="putative regulator; Not classified"
/notes="putative transcriptional regulator LysR-type;
Residues 1 to 309 of 309 are 99.67 pct identical to
residues 1 to 309 of 309 from Escherichia coli K-12 Strain
MG1655: B3243"
/codon_start=1
/transl_table=11
/product="quorum sensing Escherichia coli regulator A"
/protein_id="AAG58371.1"
/db_xref="GI:12517867"
/translation="MERLKRMSVFAKXFEFGSFTAAARQLQMSVSSISQTVSKLEDEL
QVKLANRSTRISGLTEAGRIYVQGRMLHEVDVHEOLYAFNNPTIGTLRIGCSSTM
RONVLAGTAKMLKEVGLSVNLVGTIPAPDLIADGLVDVIRVGLADDSLFSRLGA
NMVYCAKSVLTQGIPEKPADLSHSLWYSVPDNEFELIAPEGISRLYFQGRF
VNDPMLTVIRMTAGAIAYVPLMWVINEIRNGELEILLPRYQSDPRPVYALYTERDK
LPLKVQVINSITDYFVVGKLFQEMHGRGKEK"
complement (1151. .2596)
/genes="tldD"
/genes="tldD"
/genes="tldD"
complement (1151. .2596)
/function="phenotype; Not classified"
/notes="Residues 1 to 481 of 481 are 99.37 pct identical to
residues 1 to 481 of 481 from Escherichia coli K-12 Strain
MG1655: B3244"
/codon_start=1
/transl_table=11
/product="suppresses inhibitory activity of CsrA"
/protein_id="AAG58372.1"
/db_xref="GI:12517868"
/translation="MSLNLVSEQLIAANGLKHQDIFVLGQIAERRLDYGDPQSSY
HESWLEDRIRIKDGSYNIDQGVGVAISGENTGFAYADQISLLALEQSAQAARIAD
SGDGKQVOTLGAVEHSPLTYSVDPLQMSREKLDILRRVDKVAAREADKXVQEVTSLSL
GYVELIIVATDGTLLAADVRPLVRLSVLIEEDKRRERGASGGGRGFGYEPFLADLD
GEVRADAKAEAVMALVNLASAAPADGRRGVAIDDEGTGGOYNVLIEHVGLEQDFNR
GTSVFGQVCELVASLCTVDDGTWVDRGSAIDDEGTGGOYNVLIEHVGLEQDFNR
DKNLAIGMTPTNGRRRSYAHLPMPRMTNTYMLPKSTPOEIIIESVEVGIYAPNG
GGODVITSGKVFSTSEAYLIEHVGKTPKVGATLIGSGIETMQQISMGVNDLKLNDG
VVCGRGQSLPVGVCQPIKVDNLVGGTA"
complement (2752. .6552)
/genes="Z4604"
complement (2752. .6552)
/genes="Z4604"
/function="orf; Unknown function"
/notes="Residues 281 to 1266 of 1266 are 98.68 pct
identical to residues 1 to 986 of 986 from Escherichia
coli K-12 Strain MG1655: B3245"
/codon_start=1
/transl_table=11
/protein_id="AAG58373.1"
/db_xref="GI:12517869"
/translation="MRRLPGLILLTGAALVITAAALVSGRLTALPHLDARPEILNKI
ESATGPEVASQLSASQNFPLEAHDIAELKDGESFVKRVTLALDNLVQSLLENR
WQPRDTFWQLRFTNTPTITSGSDSLASHISDLFQDFHDFHDLDESEVSLTPSG
QRAELAIQTLNLDPRRAEGLVSLSTGQGYMVQVMDLDDDEGLSNGRWLQ
ADDIDLPKLMQMDNIALETAQFSLGEMTIDKGVTDGVDGVLQGSASWLGKET
HTLSVDNLTAHITRENPGQFSPDTRITMDGKPPSGALTIAWIPEQVGGKDKRS
DELIRASNLLEAGLESGIRPLAALSPALGDVWRSTQPSKINTIALDIPLOADYTR
FOASWDLAWKWLPLPRAEHFSGTSGSVENGILLTASMKQAKMPYETFRAPLEIAD
GQATISLNNDKGQDGRNDIVAKARHARGFRYLQFANDEPFLGAGISITDGS
QAWRYPPENLMDKLDVLYSGAIQGGEDNATLVYGNPQLFPYKNEQGFVFLPLR
NAXFAPQDPWALNTLGIELDINDGLMMKTGVLNGVGRASNLTAVIDYSKELLI
DADIKGKAVGYPYFDETLQSLGALTQELQDGDVNRALHLDIPNGLSVTAKGV
TLNNSLFTKPLDSTLKNLSKGFSLNGDQSEPLTASWPNQPLNDVDFSTKEGAKAYQ
VAVNLNGMQPAKTGVLPAVNEALSSVAMGKGVILPYHAGATVYVELNGDLKNV
SSHLPSPPLAPAGEPLPVNVKVDNLNSFELTQAGADHNSRWLLGQKLTLDRAIW
AASKTLPLPQSGVELMPPMNGAELALFQKGAESVGGASFPQHILTPMLLS
LGNQWNLSIVSQPTANGTQVEAQREINATLANNAPLWANIYLYNPSVAKTR
GDSTSPSPFTTERINFRGNSDAQIRCAECWFGQKFRIDSDITISGNTLTNLGI
DTGFSLTADGEMINNPGRSLKGLRGKIDAAAEFFGVTPTRQSFNVYDYLH
WRKAPQDEATLNGIHTQLGKEITTEINTGHAGOLLRLSLVDALMKRLRPFDFTF
GEGFYFDSIRSTAWIKDGMHTDITVDGLEADIAMKGSVNLVRDLNNEAVVAPIS
ATVGVAAAFVNPVGAFAASKVLGPLMSKVSILRYHISGLPDLDDPQINEVLQRPK
EKAQ"
complement (6620. .8107)
/genes="cafa"
/notes="synonym: Z4605"
complement (6620. .8107)
/genes="cafa"
/function="structural component; Cell division"
/notes="Residues 1 to 495 of 495 are 100.00 pct identical
to residues 1 to 495 of 495 from Escherichia coli K-12
Strain MG1655: B3247"
/codon_start=1
/transl_table=11
/product="bundles of cytoplasmic filaments"
/protein_id="AAG58374.1"
/db_xref="GI:12517870"
/translation="MRAKGINMTAELNVNTPSETRVAYIDGGILOEIHIERARRGIV
GNYGRVSRVLPQMGAQFVLDLKAFLHASDIMPTCEVAGEKQFTVRDISEL
VROGDLVQVVDKPLGTGKARLTDTLPSRYLPMFGASHVGSORLESERERL
KTVVAVCDQCGFIIRTAAGVGEALSDAAYLKRVTKWKRKRQTRQVLYGE
LALAQRVLRDPADELDRVDSRLTYEALLETSEYIPEMTSKLEHYGRQIPDLF
DVNEIQRALERKVELSGYLIDQTEAMTVDINTGAVGHRNDDITFNINBAT
QAIARQLRNLGGIIIDFIDNNEDHRRVRLHLSLEQALSKDRVTSYNGFSLGLV
EMPRKTRGSEIHLNCEGTCHGTQVTKVETVCYVIMREIVRVHSDRPLVYA
SPVAAALAGEESHSLAEVSEIFVGKQVQIPLNQEQFDVVMH"
complement (8079. .8672)
/genes="yhdE"
/notes="synonym: Z4606"
complement (8079. .8672)
/genes="yhdE"
/function="orf; Unknown function"
/notes="Residues 1 to 197 of 197 are 98.47 pct identical to
residues 1 to 197 of 197 from Escherichia coli K-12 Strain

```

MG1655: B3248"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical protein"
 /protein_id="AG58375.1"
 /db_xref="GI:12517871"
 /translation="MTSLYLASGSPRRQELLAQLQGVTFEIVTIGIEQRPQESAAQY
 VMLAREKARAGVIAKDLPLVIGADTVILNGEVLKPRDAEHAQAQMLRLKLSQTHQ
 VNTALADSRHVLDCLVDTVDFTFTLLDEIAGVVASGEPLDKAGAVIGIQLGICFV
 RKINGSYQAVVGLPLAVEYVYELLSENFALREKDKHG"
 complement(8681..9169)
 /gene="mred"
 /note="synonym: Z4607"
 complement(8681..9169)
 /gene="mred"
 /function="structural component; Cell envelop: Murein
 sacculus, peptidoglycan"
 /note="Residues 1 to 162 of 162 are 100.00 pct identical
 to residues 1 to 162 of 162 from Escherichia coli K-12
 Strain MG1655: B3249"
 /codon_start=1
 /transl_table=11
 /product="rod shape-determining protein"
 /protein_id="AG58376.1"
 /db_xref="GI:12517872"
 /translation="WASRSQRQWVWISFLIALLQIMPMPDNLIVFRPNWLLILL
 YWILAPHRVNVGTGVNGAILDLISGTLGVRVLAMSIIAYLALKYQLFRNLALWQ
 QALVMLLSLVVDIIVFAEFLVINVSPRPEVSVVNGVLFWFLMKVRQQA
 VQ"
 complement(9169..10272)
 /gene="mrec"
 /note="synonym: Z4609"
 complement(9169..10272)
 /gene="mrec"
 /function="structural component; Cell envelop: Murein
 sacculus, peptidoglycan"
 /note="Residues 1 to 367 of 367 are 99.45 pct identical to
 residues 1 to 367 of 367 from Escherichia coli K-12 Strain
 MG1655: B3250"
 /codon_start=1
 /transl_table=11
 /product="rod shape-determining protein"
 /protein_id="AG58377.1"
 /db_xref="GI:12517873"
 /translation="MKPIESRGPSLQIBLLAVLVLGIIADSLGTFESQIRYMDT
 AVSPFYFVSNARELLDQVSLASRDLQLENLALRQELLLKNSLLMLGQYKQENA
 RURELLGSPRLQDEQKVVQIVSTVNDPYSQVVDKGSVNGVEGQVSDKGVGQ
 VYVAKLTSRVLLICDATHALPQVLRNDIRVIAAGNCTDDLEHLPLANTDIRVD
 VLVTSGLGRFPFEGYFVAVSVSKLDQRTAVTIOARPTAGLRLLRLLLLWADRNG
 ANPMTPEEVRVANERLMQWMPVLPSDAMPKLPPEPATGIAQPTPOQPTGNAATA
 PRAPTQPAANRSLQATPPQSGAQPAPARPGQ"

Query Match 98.4%; Score 915.2; DB 1; Length 11593;
 Best Local Similarity 98.9%; Pred. No. 1.5e-226;
 Matches 920; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGGAAAGCACTAAACCGCATGTCGGTGTTCGCAAGTAGTGTGAATTTGGCTCTTTTACC 60
 DB 88 ATGGAAAGCACTAAACCGCATGTCGGTGTTCGCAAGTAGTGTGAATTTGGCTCTTTTACC 147
 QY 61 GCGCGCGCAGACAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 120
 DB 148 GCGCGCGCAGACAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 207
 QY 121 GAAGATGAGTTCAGTAAAGCTGTAAACCTAGCACAGCAGATTTGGCTGACCGAA 180
 DB 208 GAAGATGAGTTCAGTAAAGCTGTAAACCTAGCACAGCAGATTTGGCTGACCGAA 267
 QY 181 GCGGTAGAAATTACTACAGCGGCTGCCGTGCTATGCTTTCATGAAGTCAGGATGTTTCA 240
 DB 268 GCGGTAGAAATTACTACAGCGGCTGCCGTGCTATGCTTTCATGAAGTCAGGATGTTTCA 327
 QY 241 GAGCAACTGTATGCTTCAATAACACCCCGCATCGGACGCTACCGATTGGCTGTTCTTCA 300

DB 328 GAGCAACTGTATGCTTCAATAACACCCCGCATCGGACGCTACCGATTGGCTGTTCTTCA 387
 QY 301 ACTATGGCAAAATCTTCTCGCGGGCTGACAGCCAAAATGCTGAAGAATACCCAGGT 360
 DB 388 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAGAATACCCAGGT 447
 QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCCCGACCTGATTGCCGACGCTTGGAT 420
 DB 448 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCCCGACCTGATTGCCGACGCTTGGAT 507
 QY 421 GTGCGTATCGCGTCCGCGGCTGAGGATTCAGAGCTGTTTCCCGCGCTTGGCGCG 480
 DB 508 GTGCGTATCGCGTCCGCGGCTGAGGATTCAGAGCTGTTTCCCGCGCTTGGCGCG 567
 QY 481 ATGCCAATGCTGTGTCGCGCGGCGGAAAGCTATCTCACACAATACGGCATACCGGAAAA 540
 DB 568 ATGCCAATGCTGTGTCGCGCGGCGGAAAGCTATCTCACACAATACGGCATACCGGAAAA 627
 QY 541 CCGCGCGATTTAGTAGTATGATTCATGCGCTTGAATACAGCGTGGCGGCCGACATGAATTT 600
 DB 628 CCGCGCGATTTAGTAGTATGATTCATGCGCTTGAATACAGCGTGGCGGCCGACATGAATTT 687
 QY 601 GAATGATCGCACCGGAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
 DB 688 GAATGATCGCACCGGAGGATCTCGACTCGCTGATCCCGNAGGAGATTTGTGACT 747
 QY 661 AATGATCGGATGACGCTGTGCGCTGCGTGAAGCGGGTGC CGGGATCGCTACGTGCGG 720
 DB 748 AATGATCGGATGACGCTGTGCGCTGCGTGAAGCGGGTGC CGGGATCGCTACGTGCGG 807
 QY 721 CTGATGCGGATGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGTCGCGGTTAC 780
 DB 808 CTGATGCGGATGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGTCGCGGTTAT 867
 QY 781 CAGTCAGATCCACCGCGCGTTCGCTGTTATATACCGAAAAAGATAAGTCGCGCTGAAG 840
 DB 868 CAGTCAGATCCACCGCGCGTTCGCTGTTATATACCGAAAAAGATAAGTCGCGCTGAAG 927
 QY 841 GTACAGTGTGATCAACTGCTGACGATTTATTTGTTGAGTGGTAAATTTGTTTTCAG 900
 DB 928 GTACAGTGTGATCAACTGCTGACGATTTATTTGTTGAGTGGTAAATTTGTTTTCAG 987
 QY 901 GAGATGCACGGCGCGGCGGAAAGAGTAAGTAA 930
 DB 988 GAGATGCACGGCGCGGCGGAAAGAGTAAGTAA 1017

RESULT 5
 LOCUS AE016989 292497 bp DNA linear BCT 22-APR-2003
 DEFINITION Shigella flexneri 2a str. 2457T section 12 of 16 of the complete genome.
 ACCESSION AE016989
 VERSION AE016989.1
 SOURCE GI:30042681
 KEYWORDS Shigella flexneri 2a str. 2457T
 ORGANISM Shigella flexneri 2a str. 2457T
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Shigella.
 1 (bases 1 to 292497)
 Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A., Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S., Schwartz, D.C. and Blattner, F.R.
 Complete Genome Sequence and Comparative Genomics of Shigella flexneri Serotype 2a Strain 2457T
 Infect. Immun. 71 (5), 2775-2786 (2003)
 12704152
 2 (bases 1 to 292497)
 Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A., Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S.,

Schwartz, D.C. and Blattner, F.R.
 Direct Submission
 Submitted (13-JUN-2002) Genetics Laboratory, University of
 Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
 Location/Qualifiers
 1. 292497
 /organism="Shigella flexneri 2a str. 2457T"
 /mol_type="genomic DNA"
 /strain="2457T"
 /serotype="2a"
 /db_xref="taxon:198215"
 98..1169
 /gene="ygjJ"
 /locus_tag="S3326"
 /notes="frameshift"
 /pseudo
 1186..3537
 /gene="ygjK"
 /locus_tag="S3327"
 1186..3537
 /gene="ygjK"
 /locus_tag="S3327"
 /function="putative enzyme; Not classified"
 /notes="residues 1 to 783 of 783 are 95.65 pct identical to
 residues 1 to 783 of 783 from Escherichia coli K-12 ;
 B3080"
 /codon_start=1
 /transl_table=1
 /product="putative isomerase"
 /protein_id="AAP18405.1"
 /db_xref="GI:30042682"
 /translation="MKIKTLPVTCALLISPSAHANADNKNVINTGAPQYMKDY
 DYDQHRFPFDLGAWHGLLPDGTNGGPGFVALLTEYINFMASNFRLTVQD
 GKVFATLEAYIPGALVKQLTAKDVQVEMTLFATPRTSLLETKIISNKPLDLVWDQ
 ELLEKLEAGSGPLSKDTIAGEPYQKISATRDGLKVTGKVAZTWDLLTSGSEY
 OVHSLPVQTEINGNFTSKAHNGSTLTLYTSHLLTAQVSKQEMOIRDILAPAF
 YLTASQQRWEYLKGLTNPDPATPEQTRVAVKAIETLNGNRSPGAVNATVPSVT
 GRWFSQNPWDWTKQAFMAHFNPDIKENIRAVFSWQIQGDSVRPQDVGFPDL
 IAWNLSPEGGGGNNERNTPSLAASVMEVYNTQDKAWLAEMYKLVAYHDWDL
 RNRDHNGVPGATRDKAHNTSEGLFTVKGDKBEETQSGLNRYARVVEKQYDS
 LEIPAQVASESGDDAAVFGFIDKEQLDKYVANGKRSQDWTVFAENRSQDGLLG
 YSLQESVDQASVYSDNHYLAEMATILGKPEAKRYQLAQCLADYINTCFDPTTQ
 FYVDVRIEDPLANGCAGKPIVERGKPGEGSPFNAGTAQANADPVVKMLDKPEFN
 TVPLGTALTNPAGADIWRGVWVQDFWFLGKMERYGVRDDALKLATTFRAHK
 GLUTAQIOENTNPTGACQGNPFSWGAHLYMDYNDFRQ"
 3550..3632
 /notes="REP (repetitive extragenic palindromic) element,
 contains 2 REP sequences"
 3763..5781
 /gene="ygjL"
 /locus_tag="S3328"
 3763..5781
 /gene="ygjL"
 /locus_tag="S3328"
 /notes="residues 1 to 672 of 672 are 95.23 pct identical to
 residues 1 to 672 of 672 from Escherichia coli K-12 ;
 B3081"
 /codon_start=1
 /transl_table=1
 /product="putative NADPH dehydrogenase"
 /protein_id="AAP18406.1"
 /db_xref="GI:30042683"
 /translation="MSYPSLPAPLDGFTTLKNRVLMGSMHTGLEEYPDGAERLAAY
 AERAGVALIVSGGIAPDLTGVMGGMALDSQIPEHRTITEAVHQEGKIALQI
 LHTGRYSQPHLVASALQAPINRVFPHLTHEELQLDIDPARCAQARAGDQVE
 VNGSGYLNEFLTNRQSDQDQWGNRFAVEVTVRVRGNDFFIILYRLSM
 LDVLEGGFTFAEIAEAAGATIINTGIGHEARPTIATVPFGASFWYTRKL
 KGHSLPLVTNRINDPQVADDILSRGDAIVSMARPLDAELLSKASQSGRAEINT
 CIGCGOGLDQIFVGKVTSLNPRACHETKMPIPAVOKKNLAVVAGPAGLAFAIN
 AARGHONTLPFAHSEIGGOENIAKQIPCKEFVETLYAYRMIEVTGTLKLNHTVT
 AQLQAPETILLASIVRPPIIDGIDHPEVSLYDLDRDKAPGNKVAILGCGGIGF
 DANTLISQESTSNIARFNGWIGDISLQAGGLSPQGMIPKSPKQIIVMLQKAS
 KPGQGLGKTGWIHRTLLLSRGVRKMPGVSYQKIDDDGLHVINGETQVLAVDNNVIC

AGOENRALAQPLIDSGTKVHLIGGCDVAVELDARRAIAOQTRLALEI"
 complement(5826..6242)
 /gene="ygjM"
 /locus_tag="S3329"
 complement(5826..6242)
 /gene="ygjM"
 /locus_tag="S3329"
 /notes="residues 1 to 138 of 138 are 90.57 pct identical to
 residues 1 to 138 of 138 from Escherichia coli K-12 ;
 B3082"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="AAP18407.1"
 /db_xref="GI:30042684"
 /translation="MIAIDILQAGEKLTAVAPFLAGIONEQYTOALELVLDHLLND
 FENPLDLVCAKITANWEESEPAEFENMAQAMPGGIATVIRLMDQVGLTSLDLPEIG
 SKMVSRLVSGKKRLTLESHAKPLATRGISPAULFI"
 complement(6239..6553)
 /gene="ygjN"
 /locus_tag="S3330"
 complement(6239..6553)
 /gene="ygjN"
 /locus_tag="S3330"
 /notes="residues 1 to 104 of 104 are 100.00 pct identical
 to residues 1 to 104 of 104 from Escherichia coli K-12 ;
 B3083"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="AAP18408.1"
 /db_xref="GI:30042685"
 /translation="MHLITQKALKDAEKYPOHKTELVALGNTIAKGYFKKPSLKAV
 FPLSNDPKYLDKHVFNUGNELRVVAMVFFESQKCYIREVWTHKEYDFFTAVRTNG
 KK"
 6736..6822
 /notes="REP (repetitive extragenic palindromic) element,
 contains 2 REP sequences"
 complement(6838..7974)
 /gene="ygjO"
 /locus_tag="S3331"
 complement(6838..7974)
 /gene="ygjO"
 /locus_tag="S3331"
 /functions="putative enzyme; Not classified"
 /notes="residues 1 to 378 of 378 are 95.76 pct identical to
 residues 11 to 388 of 388 from Escherichia coli K-12 ;
 B3084"
 /codon_start=1
 /transl_table=1
 /product="putative enzyme"
 /protein_id="AAP18409.1"
 /db_xref="GI:30042686"
 /translation="MSHLNDFRSLTQRPFPATDDVNPLOAWEAADEYLLQQLDDTEI
 LGPVLINDAFGALSALAEHKPYSIGDSVISLATRENLINGIDESSVKFLDSTAD
 YPOQGVLLTKVPKTLALLQOLRALRKVTSITRIIAGAKARDIHTSTLEPEKVLG
 PTTTTLANKRLINCTFNEPPLADAPQVSKLEGTDTIHNHNVFSTGTLDIGAR
 FFMQHPENLEGEIVDLGCGNGVIGTLTDKNQAKVVFVDESFMASRLNVEINM
 PEALDRSEFMNNAISGVPEFRFNAVLNCPFPHQHQLTDNVAWEFMFVARRCLIKING
 ELYIVANRLHDYFHLKLFNGNCTTIATNNKFFVLKAVKLGRR"
 8023..8562
 /gene="ygjP"
 /locus_tag="S3332"
 8023..8562
 /gene="ygjP"
 /locus_tag="S3332"
 /notes="residues 1 to 179 of 179 are 98.32 pct identical to
 residues 1 to 179 of 179 from Escherichia coli K-12 ;
 B3085"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="AAP18410.1"


```

/note="Residues 1 to 309 of 309 are 99 pct identical to
residues 1 to 309 of a 309 aa protein from Escherichia
coli O157:H7 ref: NP_312143.1"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator LYSR-type"
/protein_id="AA044746.1"
/db_xref="GI:24053720"
/translation="MERUKRMSVFAKVVEFGSFTAAARQLQMSVSSISQTVSKLEDEL
QYKLLNSTRSGLTEAGRIYYQGRMLHEVQDVHEQLYAFNNTPTGTLRIGCSSTW
AGNLNFTAKMLKEYPGLSANLVGTI PAPDLADGLDVIWIRVAGALQSSLSFRRLGA
MNVVCAKSYLTQYGIPEKPADLSHSLWYSVRPDNEFELIAPEGISTRLIPQGRF
VTNDPMTLVRLTACAGIAYVPLMWVINEINRGELEILLPRYQSDPRPVVLYTEKDK
LPLKVQVINSLTDFEVGKLPQEMHGRGKEK"
complement(1285..2730)
/locus_tag="SF3283"
/locus_tag="SF3283"
complement(1285..2730)
/locus_tag="SF3283"
/locus_tag="SF3283"
/note="Residues 1 to 481 of 481 are 100 pct identical to
residues 1 to 481 of a 481 aa protein from Escherichia
coli K12 ref: NP_417711.1"
/codon_start=1
/transl_table=11
/product="suppresses inhibitory activity of CsrA"
/protein_id="AA044747.1"
/db_xref="GI:24053721"
/translation="MSLNLVSEQLLAANGLKHQDLPAILQGLAERLLDGYLQSSY
HESWLEDRIRKDSYNDQGVRAISGEKTFAYADQISLLALEQSAQAARTVRD
SGDGVQVLAFTSVDPQLQMSREKLDLRVDKVARADKRVGEVTAASL
GYVELLVAATDGTLLAADVPLVRLSVSVLEEDKRGKRGASGGGGRGFEFFLADLD
GEVRLDAWAKEAVALNVLNLSAAPAAGTVPVLCAGWPGLLHEAVGHGLEGDFNRR
GTVSPSGVGVASELCTVVDGTMVDRGSAVDDEGTGQGVNVLLENGILKGYMQ
DKLNRALMGMTGTNGRRESYAHLEPMRNTVYMLPGKSTPQEIIESVEYGIYAPNFG
GQGVDTLTKGFTSSEAYLIENGKVTAPVKGATLIGSGIETMQQISVGNGLKNDG
VGVCGKGGCLPVGQGTLPKVDNLTVGGTA"
complement(2886..6686)
/locus_tag="SF3284"
/locus_tag="SF3284"
complement(2886..6686)
/locus_tag="SF3284"
/locus_tag="SF3284"
/note="Residues 1 to 1266 of 1266 are 99 pct identical to
residues 1 to 1266 of a 1266 aa protein from Escherichia
coli K12 ref: NP_417712.1"
/codon_start=1
/transl_table=11
/product="uncharacterized membrane protein"
/protein_id="AA044748.1"
/db_xref="GI:24053722"
/translation="WRRLLPGILLTGAALVIAALLVSLRIALPHLNAMRPEILNKI
ESATGMPVEASQSWQNFPTLEAHDIRAEIKDGFYSVKRVTALDVWQSLHMR
WQFRLTFWQLEFRNTPTITSGSDSDSLEASHISDLRQPDHDLRSEVSLTPSG
QRAELAIQTLWLNPRBRARGLVLSLQGHGVQVMDLRDDEGLLNGRVWLQ
ADDIDLKPMQMNQNTALETAQSLSGWMTIDKGVTDGVDWMLQGGASWLGKQT
HTLSDNLTATHITRENQWQFSIPDTRITMDGQWPSGALTAMIPEQDVGKONKRS
DELRASNLLEAGLEGRPLAAKLSPALGDVWVRSTQPSKINTLALDIPQAADKTR
FOASNDLAWKQKLLPGAHEPSGLTSGSVENGILLTASMKQAKMPEYTVFRAPLEAD
GOATISLWNLKGFOLDGRIIDVAKAVHARGFPYLOPANDPEMLGLAGISTDDGS
QARVPEPNMLKGLDVLYLSGAIQGEADNATLVYGGNPQLPYKHNEGQFELVPLR
NAKFAQPDWPAITNLDLHDFINDGLWKNKTDGCVNLGVRASNLTAVIDPISKEKLLI
DADIKPGKAVGPFYDETPKDSLGATQELQDGDVNARLHLDIPLNGELVITKAGEV
TURNSLFTKPLDLSIKNLSGKFSFINDGLQSEPTASFNQPLNDFSTKEGAYQY
VAVNLGNWQPAKTGYLPEAVNEALSVDGKVGIDLPYHAGATYVVELNGLNWY
SGLHPSPLKPAGEPLAVNVDGNSLNGSFELTGOAGADNHFNSRLWLLQKLTLDRAIV
AADSKTLPPEQSGVELNMMPMNGEALWALPQKAARSVGAASFPOHITLRTPLMS
LQGNQWNLISVQSPGTANLMLVAGGRINATLAVNNAPLANIKLYNPSVAKTR
GSTSSPPPTTERINFRGNWPAQIRCTECNFWGQKFRISDITISGDTLITNLGI
DTGFRSLTADGGEVNNPGRNERSLKGURQKIDAAEFVGTPIRQSSFNVDYDLH
WRKAPQWQDEATLNGIITHLQKGEITINTGACQLLELLSVDMKMKIRDFDRF
GEGFYFDSIRSTAWIKGWHTDDTLVDGLEADIAKGSVNLVRDLNNEAVAPEIS
ATVGVAAAFANVPGAAVFAASKVLGLPWSKVSILRYHISGPLDDPDQINEVLRQPK

```

```

EKAQ"
complement(6754..8241)
/locus_tag="SF3285"
/locus_tag="SF3285"
complement(6754..8241)
/locus_tag="SF3285"
/locus_tag="SF3285"
/note="Residues 1 to 495 of 495 are 99 pct identical to
residues 1 to 495 of a 495 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289814.1"
/codon_start=1
/transl_table=11
/product="bundles of cytoplasmic filaments"
/protein_id="AA044749.1"
/db_xref="GI:24053723"
/translation="MRKGINMTAELLVWVTPSTRVAYIDGILQBIHIERARRGIV
GNYKGRVSRVLPQMAAFVDIGLDAKAFHASDIMPHTECVAGEQKQFTVRDISEL
KVVQDLVQVWVDPLGTGARTLITDITLPSRYLVPFGASHVGVQRIESESEREL
KQVAEYCDQGGFIRTAAGVGGEALASDAAYLKRVMTKVMERKKRPTQRYLYGE
LALAORVLRDPAEALDRVDSRLTVEALLETSEVYIPMTSKLEHYTGROPIFLF
DVNEIQORALERKVELKSGYLIIDOTEAMTVDINTGAFVGHENLDDTIFNINIEAT
QAIARQRLNGLGIIIDFIDNNNDHRRVRLHLSLEQALSKDRVKTSTVNGPSALGIV
EMTRKRTRESIEHVLNCEPTCHGRGTGVTKEVTCYEIMREIVRVVHHVHDSDFLVIA
SPAAEALKGEEHSLAEVEIFVGQKVQIEPLYNQEQPDVVMN"
complement(8213..8806)
/locus_tag="SF3286"
/locus_tag="SF3286"
complement(8213..8806)
/locus_tag="SF3286"
/locus_tag="SF3286"
/note="Residues 1 to 197 of 197 are 97 pct identical to
residues 1 to 217 of a 217 aa protein from Escherichia
coli gb: AA58051.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA044750.1"
/db_xref="GI:24053724"
/translation="MTSLYLAGSPRRQELLAQLGVTFERIVTGIEEQPOBSAQOY
VVRAREKAQAGVATQADPLVGLADIIVTLNGEVLKPRDAEHAQMRLKLSGQTHO
VMTAVALASCHILDCLVVTDVTFRLTDEDIAGYVAGSEPLDKAGAYGQGLGGCFV
RKNGSYHAVVGLPLVETVELLSNPNALREKDKHDG"
complement(8815..9303)
/locus_tag="SF3287"
/locus_tag="SF3287"
complement(8815..9303)
/locus_tag="SF3287"
/locus_tag="SF3287"
/note="Residues 1 to 162 of 162 are 98 pct identical to
residues 1 to 162 of a 162 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289816.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA044751.1"
/db_xref="GI:24053725"
/translation="MASYRSQGRWVWLSFLIALLIQIMPDPNLVFRPNWVLLIL
WYIALLPFRVGTGFMGAILDLISGLVGVWLVMSIAYLVALKYQLFRNLALWQ
QALVWMLLSIVDIIIVFAEFLVINTVSPRPEVFWSSVWVGLWPMWIFLLMRKVRQOFA
VQ"
complement(9303..10406)
/locus_tag="SF3288"
/locus_tag="SF3288"
complement(9303..10406)
/locus_tag="SF3288"
/locus_tag="SF3288"
/note="Residues 1 to 367 of 367 are 99 pct identical to
residues 1 to 367 of a 367 aa protein from Escherichia
coli K12 ref: NP_417716.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA044752.1"

```

```

/db_xref="GI:24053726"
/translation="MKPIFSRQPSQIRLILAVLVALGIIIDSLRGTFISOIRTYMDT
AVSPFYVSNARELLDGIQSLARLDQLEENLRALQELLKNSLEMLGQYKQENA
RURELLGSLRQDQWMTQVISTVNDPFSDDQVLDKGSVNGVIEGQVPSIDKGVGQ
VVAVALKTSRVLLI CDATHALPVLKLRNDIRVIAAGNCTDDQLHLEHLPANTDIRVG
VLVTSGLGRFPEVAVSVSKLDRAYTIOARPTAGLRLRYLLLLWADNRNG
ANMPTPEVHRVAMERLMPQVLPSPDAMPKLPPEPATGIAOPTPOCPATGNAATA
PVATQPAANRSPORATPQSGAOPPARAPGQ"
complement (10472..11590)
/locus_tag="SF3289"
/locus_tag="c3797"

Query Match      98.1%; Score 912.4; DB 1; Length 11734;
Best Local Similarity 98.8%; Pred. No. 7.8e-226;
Matches 919; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGGACGACTAAACCGCATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 60
DB 228 ATGGAACGACTAAACCGCATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 287
QY 61 GCCGCGCCGACAGCAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGCTATCAAACTG 120
DB 288 GCCGCGCCGACAGCAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGCTATCAAACTG 347
QY 121 GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCAGCAGGATTTGGCTGACCGAA 180
DB 348 GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCAGCAGGATTTGGCTGACCGAA 407
QY 181 GCCGTPAGATTTTACTACGAGGCTCCCTGCTGATGTTTCATGAAGTCAGAGATGTTTCA 240
DB 408 GCCGTPAGATTTTACTACGAGGCTCCCTGCTGATGTTTCATGAAGTCAGAGATGTTTCA 467
QY 241 GAGCAACTGTATGCTTCAATAACCCCATCGGAGCGCTACGATGCGTGTGTTCTTCA 300
DB 468 GAGCAACTGTATGCTTCAATAACCCCATCGGAGCGCTACGATGCGTGTGTTCTTCA 527
QY 301 ACTATGGCACAAATGTTCTCCGCGGCTGACAGCCAAATCTCAAGAAATACCCAGGT 360
DB 528 ACTATGGCACAAATGTTCTCCGCGGCTGACAGCCAAATCTCAAGAAATACCCAGGT 587
QY 361 TTGAGCGTCAATGCTGTTACCGGAATTCAGACCCCGGACCTGATTCGCGAGCGTCTGAT 420
DB 588 TTGAGCGCAATCTGTTACCGGAATTCAGACCCCGGACCTGATTCGCGAGCGTCTGAT 647
QY 421 GTGGTATTCGGGTGCGCGCTGACGAGTTCAGGCTGTTTCCGCGGCTGCGCGCG 480
DB 648 GTGGTATTCGGGTGCGCGCTGACGAGTTCAGGCTGTTTCCGCGGCTGCGCGCG 707
QY 481 ATGCCAATGTTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
DB 708 ATGCCAATGTTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 767
QY 541 CCGCGCGATTTGAGTAGTCAATTCATGCTTGAATACAGGTCGCGCGCGGACAAATTT 600
DB 768 CCGCGCGATTTAAGTAGCAATTCATGCTTGAATACAGGTCGCGCGCGGACAAATTT 827
QY 601 GAAGTATCGACCGGAGGAGTCTCGACTCCGCTGATCCCAAGAGGAGATTTGTGACT 660
DB 828 GAAGTATTTGACCGGAGGAGTCTCGACTCCGCTGATCCCAAGAGGAGATTTGTGACT 887
QY 661 AATGATCCGATGACGCTGTCGCTGGTGTGACGCGGCGGTGCGGATCGCCTACGTCGCG 720
DB 888 AATGATCCGATGACGCTGTCGCTGGTGTGACGCGGCGGTGCGGATCGCCTACGTCGCG 947
QY 721 CTGATGCGGTGATCAACGAGATCAATCTGCGGAGCTGGAGATCCTCTCGCGGTTAC 780
DB 948 CTGATGCGGTGATCAACGAGATCAATCTGCGGAGCTGGAGATCCTCTCGCGGTTAT 1007
QY 781 CAGTCAGATCCAGCGCGGCTTTATCGGTTATATACCGAAAAAGATAAGCTCGCGTGAAG 840
DB 1008 CAGTCAGATCCAGCGCGGCTTTATCGGTTATATACCGAAAAAGATAAGCTCGCGTGAAG 1067
QY 841 GTACAGGTCGTGATCAACTCGCTGACGAGTATTTGTTGAGTCCGTAATTTGTTTCAG 900

```

```

Db 1068 GTACAGTCTGATCAACTCGCTGACCGATATTTGTTGAGTCCGTAATTTGTTTCAG 1127
QY 901 GAGATCAGCGGCGCGGCGGAGAGAGATTA 930
Db 1128 GAGATCAGCGGCGCGGCGGAGAGAGATTA 1157

RESULT 7
LOCUS AE016767 300359 bp DNA linear BCT 24-DEC-2002
DEFINITION Escherichia coli CFT073 section 13 of 18 of the complete genome.
ACCESSION AE016767 AE014075
VERSION AE016767.1 GI:26110054
KEYWORDS
SOURCE Escherichia coli CFT073
ORGANISM Escherichia coli CFT073
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 300359)
AUTHORS Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence
of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)
12471157
REFERENCE 2 (bases 1 to 300359)
AUTHORS Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
1. 300359
/organism="Escherichia coli CFT073"
/mol_type="genomic DNA"
/strain="CFT073"
/db_xref="taxon:199310"
188..547
/genes="yqiC"
/locus_tag="c3796"
188..547
/genes="yqiC"
/locus_tag="c3796"
188..547
/genes="yqiC"
/locus_tag="c3796"
/notes="Escherichia coli K-12 ortholog: b3042; Escherichia
coli O157:H7 ortholog: z4400"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiC"
/protein_id="AAN82241.1"
/db_xref="GI:26110055"
/transl_table=11
/translation="MTYPTTVALNTYQYRETTWIDPKKIEQIAQVHESMPKQIRE
FEDVKKTRQTLQALRLDLVSRREFDVQTVLLRTREKLALLQRISLEARYNP
VAELQSPPAIPVDKAE"
complement (582..788)
/genes="gls"
/locus_tag="c3797"
188..788
/locus_tag="c3797"
complement (582..788)
/genes="gls"
/locus_tag="c3797"
188..788
/locus_tag="c3797"
/function="putative enzyme: Macromolecule synthesis,
modification: Polysaccharides - (cytoplasmic)"
/notes="Escherichia coli K-12 ortholog: b3049; Escherichia
coli O157:H7 ortholog: z4401"
/codon_start=1
/transl_table=11
/product="Glycogen synthesis protein gls"
/protein_id="AAN82242.1"
/db_xref="GI:26110056"

```



```
/translation="MLMDHSLNSLNNFDFPLARSFARMAHGRPVDILAVTGNMDEHR
TWFCARYAWYCOQMMQTRLELEH"
gene
1052..1681
/genes="yqiJ"
/locus_tag="c3798"
1052..1681
/genes="yqiJ"
/locus_tag="c3798"
/translation="putative"
/function="putative"
/notes="Escherichia coli K-12 ortholog: b3050; Escherichia
coli O157:H7 ortholog: z4402"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiJ"
/protein_id="AAN82243.1"
/db_xref="GI:26110057"
/translation="MLFADYNTPYLFAISFVLLIGLLEIFALICGMLSGALDAHLD
HVSITIGHISQALHINIGRLPALVVLCLLAGRGLIGILLQACVTLWSPPLNLF
VVPVSLFTIIAHTGKVAPWIPRDHSSAITEYVGSMLITGHQATSGNCEGK
LDQCGQIHLLEFEFGTKFGDKVLLICLSAIRYLAENPNWSQIL"
1708..3369
/genes="yqiK"
/locus_tag="c3799"
1708..3369
/genes="yqiK"
/locus_tag="c3799"
/function="putative membrane"
/notes="Escherichia coli K-12 ortholog: b3051; Escherichia
coli O157:H7 ortholog: z4403"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiK"
/protein_id="AAN82244.1"
/db_xref="GI:26110058"
/translation="WDDIVNSVPSMFTAIIVACILFIIGIFARLYRSASAEQAFVR
TGLSQKVMGSGAIVMFIPIHEIIBINNTLKVLSRSTIDSLITKDRMYDVVVAFF
VRKPSVEGIAAQTGLQRTLSPEDRLMVEDKFDALRATAQMTHELDQTRNF
VGVONTVAEDLSKNGLESVLSNFQTSKEHFNPNNAFDAEGLTKLTOTERRRR
ERNEQDVAVRSKNDALSRKLEIEQGEAFVLEQEQVTFEATONAKIAFEA
ERRAEQTRILAEIQIQTETEDREQAVRSKVEAEVRVKEIEQQQVTEIANQTKS
IAIAKQEQSQAERANALAEVSAQONVETTRQTAEDRAKQVAILAAQAEIK
AVELTVRAKEAEMQAAIVELAEATRKKGLAEAAQALNDAINVLSDQTSK
FKIALQALPAVIEKSPFMKIDIGIKIIVDGLNRGGATGDANTGNVGGNLAELQ
SAALSVRTQAPILDSILNEIGVSGSLAALTSSLSSTTPVEKAE"
complement(3788..5221)
/locus_tag="c3800"
complement(3788..5221)
/locus_tag="c3800"
/EC_number="2.7.-.-"
/function="putative"
/notes="Escherichia coli K-12 ortholog: b3052; Escherichia
coli O157:H7 ortholog: z4405"
/codon_start=1
/transl_table=11
/product="ADP-heptose synthase"
/protein_id="AAN82245.1"
/db_xref="GI:26110059"
/translation="MKVTLPEFERAGVVMGVMDLRVYVGYTSRISPEAPVPVVKVN
TTEERPAGAANVANIASIGANARLVGLTIDDAARALKSLADVNVKCDVSVPTHP
TITKRLVSRNQVLLRDPFEGFEGVDQPLHERINQALSIGALVLSDYAKGALASV
QMIIQARAGVFLIDPKGTDFERYRGATLLITPNLSFEAVVWCKTEEBEIVRGMK
LADVELSALLYTRSEQNSLQPGKAPLHMTQAEYVDYTGAGDTVIGVIAATLAA
GNSLEACFFFAAAGAVVVGKLTSTVSPILENAVRGRADTFCGVTMEELKLVAA
ARKRGKVTMTNGVFDILHAGHVSYLARAKLGDRLLIVAVNSDASTKELKGDSPVNP
LEORMIVLGAEVAVDWSFEEDTPORLIATLIPLLLVKGGDYKFEETAGSKYEWANG
GEVLVNFEDGCGSTNLIKIQLDKKG"
complement(5269..8220)
/genes="glnE"
/locus_tag="c3801"
complement(5269..8220)
/genes="glnE"
/locus_tag="c3801"
/EC_number="2.7.7.42"
```

```
/function="enzyme; Proteins - translation and
modification"
/notes="Escherichia coli K-12 ortholog: b3053; Escherichia
coli O157:H7 ortholog: z4406"
/codon_start=1
/transl_table=11
/product="Glutamate-ammonia-ligase adenyllyltransferase"
/protein_id="AAN82246.1"
/db_xref="GI:26110060"
/translation="MRPGNALKLNISVMQTIIRNRSGCTAENVNOARIKSLMKPLSSP
LQOYVOTVVERLPEPLAEESLSAQAKSVLTFFSDVDSVIAHPWEMTELESOPQADE
KQYVASHLQEALSNVSDERAGLMRELRLFRIRINVRITAKATLALVTEESLQOLSHLA
ETLIVARWLYDACCREWGTPCNAQEAPLLILNGKLGGLGELNPFSSDIDIFAWP
EHGTCGGRRELDNAQFTFRMQRLIKVLDQTDQGVFVYVDMRLRPFSGSGPLVLSF
AALEDYYEQQRDMERYAMKARIMGSDGVYANERLAMLRFVFRYIDFVIOQLR
NMKGMTAREVRRRLTDNIKLAGGIREIEFIQVQFLIRGGREPSLQSRALLPTLSA
IAALHLSANDAEQLAVLYFLNRELIGDDESETQEEESLSEOMRELQDALQEDTTPVLAH
WPOLTVLTVANANVRVFNELIGDDESETQEEESLSEOMRELQDALQEDTTPVLAH
LSDDRAQVLMILIADEFKELDKETIGPRGQVLDHMLPHLLSDVCAREDAVTLRIT
ALIVIVIRITYLLELSEFFPAUKHLISCAAFMIAQASQARIPLLIDELLDNTYLIQ
PTATDAYDELRYLLRPEDDEEQLEALRFQKQALLRAIAAADAGTLPVMKVSDD
LTLWAEAMIDAVVQAWVMVARYGNPKHNEREGRFVAVVGGKLGWELRPSGDL
LIPLHCPMDAMTDGSEIDGQFYLRLAQRIIMHLEFSTRSSGILYEVDAIRLPSGAA
GMLVTSAAEFADYQKNEATWHEQALVARVYVGDPLQTAHFDVAREIMTLPREGKT
LOTEVREMEKMBHLNKHREDPDIKADGGITDIEFTQYLVLYVAHEKPKLTRKS
DNVRIELLLQNDIMEEQEAMALTRAYTTLRDELHLLALQELFCHVSEDCFTAARELV
RASQKWLVEE"
complement(8132..9460)
/genes="ygiF"
/locus_tag="c3802"
complement(8132..9460)
/genes="ygiF"
/locus_tag="c3802"
/notes="ORFXE; Escherichia coli K-12 ortholog: b3054;
Escherichia coli O157:H7 ortholog: z4407"
/codon_start=1
/transl_table=11
/product="Hypothetical protein ygiF"
/protein_id="AAN82247.1"
/db_xref="GI:26110061"
/translation="MGVPSGGMAQEIEIKFIVNHSVAVALRDHNLNTLGEHHPDVQ
LNNIYETPNWLRHDMGLRIRGCRYEMTMKVAGRTVGGHLQPEYNVALSEPTL
DLNLTVEWPNGLPADLASVQPLPFDYFYEKWLVEVDGSRISIAFPQGVKAGE
FAPPICELEBELLSDMRAVLKLANVLSQTLRGQSLSKAARGYHIAQONPARETKP
TTLHVAAKADVEQGLEALELAQWYHEELWVGNDAAKEQVLAAGLVRHTLML
FGGIVPRKASTHLRDLITQCEATIASAVSAVTAVYSTETAMAKLALTENLWSKWQPF
LDKAKSKSDSFRFADILHSRAAEKLSVFCQPLGDRYDRDQPLRLTRIDDSILLIA
GYDVPVVAQWLENWQGLRHAIATQRIEIEHFRNEANNQEPFWLHSGKR"
complement(9503..9955)
/locus_tag="c3803"
complement(9503..9955)
/locus_tag="c3803"
/function="putative enzyme; Transposases"
/notes="Escherichia coli O157:H7 ortholog: z5815"
/codon_start=1
/transl_table=11
/product="putative Transposase"
/protein_id="AAN82248.1"
/db_xref="GI:26110062"
/translation="MWLFPMYNIISTFFWQYVCEINRSRYSAYLHVVHVPMKYVR
FVFGELHANMQRYVAEVCADFGAEKCEGDADHVMHLEIYPLVQLTKLVNLSKSV
TSRLRNDFTDLRAAYSKPVLWSRYFISCGGAPLEVVKYIQNQRG"
9968..11173
/genes="ydcM"
Query Match 94.8%; Score 882; DB 1; Length 300359;
Best local similarity 96.8%; Pred. No. 7.7e-218;
Matches 900; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 ATGGAACGACTAAACGCGATGTCGGTGTGTTGCCAAAGTAGTGAATTGGCTCTTTTACC 60
|||||
Db 182614 ATGGAACGACTAAACGCGATGTCGGTGTGTTGCCAAAGTAGTGAATTGGCTCTTTTACC 182673
```



```
QY 61 GCGCGCCGACAGCTACAGATGAGCGTTTCCTCATCAGTCAGACGGTATCAAACTG 120
Db 182674 GCGCGCCGACAGCTACAGATGAGCGTTTCCTCATCAGTCAGACGGTATCAAACTG 182733
QY 121 GAAGATGAGTTCAGGTAAGCTGTTAAACCTGATGACGACGATGGCTGACCAA 180
Db 182734 GAAGATGAGTTCAGGTAAGCTGTTAAACCTGATGACGACGATGGCTGACCAA 182793
QY 181 GCGGTAGATTTACTACAGGCGTCCGCTGCTGCTTCTATGAAGTACAGATGTTTCA 240
Db 182794 GCGGTAGATTTACTACAGGCGTCCGCTGCTGCTTCTATGAAGTACAGATGTTTCA 182853
QY 241 GAGCACTGATGCTTCAATACACCCCTGCGGACGCTGAGTGGCTGTTCTTCA 300
Db 182854 GAGCACTGATGCTTCAATACACCCCTGCGGACGCTGAGTGGCTGTTCTTCA 182913
QY 301 ACTATGCGCAAAATGTTCTCGCGGCTGACGACCAAAATGCTGAAAGATACCCAGT 360
Db 182914 ACTATGCGCAAAATGTTCTCGCGGCTGACGACCAAAATGCTGAAAGATACCCAGT 182973
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGGACCTGATGCGGACGTTGAT 420
Db 182974 TTGAGTGTCAATCTGTTACCGGAATTCAGCCCGGACCTGATGCGGACGTTGAT 183033
QY 421 GTGGTATCGCGTGGCGGCTGAGGATTCAGGATTCAGGCTGTTTCCGCGCTGGCGCG 480
Db 183034 GTGGTATCGCGTGGCGGCTGAGGATTCAGGATTCAGGCTGTTTCCGCGCTGGCGCG 183093
QY 481 ATGCCAATGTTGTTGCGCGCGAAAGTCTCTCACAAATACGGATACCGGAAAA 540
Db 183094 ATGCCAATGTTGTTGCGCGCGAAAGTCTCTCACAAATACGGATACCGGAAAA 183153
QY 541 CCGCGGATTTGAGTGTGATTCATGCTGCTGTTGATGAGGCTGCGGCGGACGATTT 600
Db 183154 CCGCGGATTTGAGTGTGATTCATGCTGCTGTTGATGAGGCTGCGGCGGACGATTT 183213
QY 601 GAAGTATCGCACCGGAGGATCTGACTCGCTGATCCCAAGGAGATTTGCTGACT 660
Db 183214 GAAGTATCGCACCGGAGGATCTGACTCGCTGATCCCAAGGAGATTTGCTGACT 183273
QY 661 AATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 183274 AATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183333
QY 721 CTGATGCTGCTGATCAACAGATCAATCGTGGGAGCTGAGATCCTGCTGCGGTTAC 780
Db 183334 CTGATGCTGCTGATCAACAGATCAATCGTGGGAGCTGAGATCCTGCTGCGGTTAC 183393
QY 781 CAGTCAGATCCAGCGCGCTTTATCGTTATATACGAAAGATGATGCTGCGCTGAAG 840
Db 183394 CAGTCAGATCCAGCGCGCTTTATCGTTATATACGAAAGATGATGCTGCGCTGAAG 183453
QY 841 GTACAGTCTGATCAACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 183454 GTACAGTCTGATCAACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183513
QY 901 GAGATCAGCGCGCGGGAAGAGAGTAA 930
Db 183514 GAGATCAGCGCGCGGGAAGAGAGTAA 183543
```

```
RESULT 8
AC020970
LOCUS AC020970 212936 bp DNA linear HTG 10-FEB-2000
DEFINITION Mus musculus clone RP23-252M21, WORKING DRAFT SEQUENCE, 144
unordered pieces.
ACCESSION AC020970
VERSION AC020970.1 GI:6691260
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 212936)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 212936)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>
-----Summary Statistics
Estimated insert size: 212936; sum-of-contigs estimation
Estimated insert size: 157000; pulse field gel estimation
Quality coverage: 4.20x in Q20 bases; pulse field gel estimation
Quality coverage: 3.10x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 144 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1159: contig of 1159 bp in length
* gap of unknown length
* 1160 2363: contig of 1204 bp in length
* gap of unknown length
* 2364 3604: contig of 1241 bp in length
* gap of unknown length
* 3605 5188: contig of 1584 bp in length
* gap of unknown length
* 5189 6372: contig of 1184 bp in length
* gap of unknown length
* 6373 7737: contig of 1365 bp in length
* gap of unknown length
* 7738 9783: contig of 2046 bp in length
* gap of unknown length
* 9784 11046: contig of 1263 bp in length
* gap of unknown length
* 11047 12223: contig of 1177 bp in length
* gap of unknown length
* 12224 13492: contig of 1269 bp in length
* gap of unknown length
* 13493 14669: contig of 1177 bp in length
* gap of unknown length
* 14670 15886: contig of 1217 bp in length
* gap of unknown length
* 15887 17089: contig of 1203 bp in length
* gap of unknown length
* 17090 18308: contig of 1219 bp in length
* gap of unknown length
* 18309 19518: contig of 1210 bp in length
* gap of unknown length
* 19519 20710: contig of 1192 bp in length
* gap of unknown length
* 20711 21932: contig of 1222 bp in length
* gap of unknown length
* 21933 23606: contig of 1674 bp in length
* gap of unknown length
* 23607 25328: contig of 1722 bp in length
* gap of unknown length
* 25329 26649: contig of 1321 bp in length
* gap of unknown length
* 26650 27849: contig of 1200 bp in length
* gap of unknown length
* 27850 29135: contig of 1286 bp in length
* gap of unknown length
* 29136 30700: contig of 1565 bp in length
* gap of unknown length

* 30701 31877: contig of 1177 bp in length
* gap of unknown length
* 31878 33065: contig of 1188 bp in length
* gap of unknown length
* 33066 34267: contig of 1202 bp in length
* gap of unknown length
* 34268 35474: contig of 1207 bp in length
* gap of unknown length
* 35475 36650: contig of 1176 bp in length
* gap of unknown length
* 36651 37859: contig of 1209 bp in length
* gap of unknown length
* 37860 39072: contig of 1213 bp in length
* gap of unknown length
* 39073 40406: contig of 1334 bp in length
* gap of unknown length
* 40407 41630: contig of 1224 bp in length
* gap of unknown length
* 41631 42864: contig of 1234 bp in length
* gap of unknown length
* 42865 44200: contig of 1336 bp in length
* gap of unknown length
* 44201 46087: contig of 1887 bp in length
* gap of unknown length
* 46088 48227: contig of 2140 bp in length
* gap of unknown length
* 48228 49518: contig of 1291 bp in length
* gap of unknown length
* 49519 50910: contig of 1392 bp in length
* gap of unknown length
* 50911 52420: contig of 1510 bp in length
* gap of unknown length
* 52421 53613: contig of 1193 bp in length
* gap of unknown length
* 53614 55004: contig of 1391 bp in length
* gap of unknown length
* 55005 56300: contig of 1296 bp in length
* gap of unknown length
* 56301 57505: contig of 1205 bp in length
* gap of unknown length
* 57506 58686: contig of 1181 bp in length
* gap of unknown length
* 58687 60407: contig of 1721 bp in length
* gap of unknown length
* 60408 61632: contig of 1225 bp in length
* gap of unknown length
* 61633 62880: contig of 1248 bp in length
* gap of unknown length
* 62881 64220: contig of 1340 bp in length
* gap of unknown length
* 64221 65420: contig of 1200 bp in length
* gap of unknown length
* 65421 66605: contig of 1185 bp in length
* gap of unknown length
* 66606 68682: contig of 2077 bp in length
* gap of unknown length
* 68683 70712: contig of 2030 bp in length
* gap of unknown length
* 70713 71895: contig of 1183 bp in length
* gap of unknown length
* 71896 73462: contig of 1567 bp in length
* gap of unknown length
* 73463 74806: contig of 1344 bp in length
* gap of unknown length
* 74807 76205: contig of 1399 bp in length
* gap of unknown length
* 76206 77580: contig of 1375 bp in length
* gap of unknown length
* 77581 79081: contig of 1501 bp in length
* gap of unknown length
* 79082 80314: contig of 1233 bp in length
* gap of unknown length
* 80315 81899: contig of 1585 bp in length

* 81900 gap of unknown length
* 83140: contig of 1241 bp in length
* gap of unknown length
* 83141 84665: contig of 1525 bp in length
* gap of unknown length
* 84666 85888: contig of 1223 bp in length
* gap of unknown length
* 85889 87155: contig of 1267 bp in length
* gap of unknown length
* 87156 88435: contig of 1280 bp in length
* gap of unknown length
* 88436 89955: contig of 1520 bp in length
* gap of unknown length
* 89956 91277: contig of 1322 bp in length
* gap of unknown length
* 91278 92513: contig of 1236 bp in length
* gap of unknown length
* 92514 93917: contig of 1404 bp in length
* gap of unknown length
* 93918 95448: contig of 1531 bp in length
* gap of unknown length
* 95449 96800: contig of 1352 bp in length
* gap of unknown length
* 96801 97978: contig of 1178 bp in length
* gap of unknown length
* 97979 99179: contig of 1201 bp in length
* gap of unknown length
* 99180 100437: contig of 1258 bp in length
* gap of unknown length
* 100438 101827: contig of 1390 bp in length
* gap of unknown length
* 101828 103117: contig of 1290 bp in length
* gap of unknown length
* 103118 104550: contig of 1433 bp in length
* gap of unknown length
* 104551 105854: contig of 1304 bp in length
* gap of unknown length
* 105855 107523: contig of 1669 bp in length
* gap of unknown length
* 107524 109208: contig of 1685 bp in length
* gap of unknown length
* 109209 110430: contig of 1222 bp in length
* gap of unknown length
* 110431 111662: contig of 1232 bp in length

Query Match 72.6%; Score 675; DB 2; Length 212936;
Best Local Similarity 98.1%; Pred. No. 4.4e-164; Indels 17; Gaps 17;
Matches 862; Conservative 0; Mismatches 0

Qy 52 TCTTTTACCGCCCGCCGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTA 111
Db 168976 TCTTTTACCGCCCGCCGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTA 169034

Qy 112 TCAAACTGGAAGATGAGTGCAGGTAAAGCTGTTAAACGCTAGCAGCAGCAGCATGGC 171
Db 169035 TCAAACTGGAAGATGAGTGCAGGTAAAGCTGTTAAACGCTAGCAGCAGCAGCATGGC 169093

Qy 172 CTGACCGAAGCCGCTAGAAATTTACTACAGGGGTCGCCGTGCTATGCTTCATGAAGTCAG 231
Db 169094 CTGACCGAAGCCGCTAGAAATTTACTACCA-GGCTGCCGTGCTATGCTTCATGAAGTCAG 169152

Qy 232 GATGTTCTAGAGCACTGTATGCTTCAATATACACCCCATCGGGAGCGTACGCAATGGC 291
Db 169153 GATGTTCTAGAGCACTGT-TGCCTTCAATATACACCCCATCGGGAGCGTACGCAATGGC 169211

Qy 292 TGTTCCTTCAACTATGGCACAATAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAGAA 351
Db 169212 TGTTCCTTCACTATGGCACAATAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAG-A 169269

Qy 352 TACCCAGGTTGAGCGTCAATCTGGTTACCGGAATTCAGGCCCCGACCTGATTCGCGAC 411
Db 169270 TACCCAGGTTGAGCGTCAATCTGGTTACCGGAATTCAGGCCCCGACCTGATTCGCGAC 169328

QY	412	GCTCTGGATGCTGATCGCGTCCGCGCTGCAGGATTCACGCTGTTTCCGCGCT	471
Db	169329	GCTCTGGATGCTGATCGCGTCCGCGCTGCAGGATTCACGCTGTTTCCGCGCT	169387
QY	472	CTGGCGCGCATCCCAATGCTGCTGCGCGCGGAAAGCTATCTCACAAATACGGCATA	531
Db	169388	CTGGCGCGCATCCCAATGCTGCTGCGCGCGGAAAGCTATCTCACAAATACGGCATA	169446
QY	532	CCG3AAAACCCCGCGCATTTGAGTAGTATTCATGCTTGAATACAGGCTCGGCGCGAC	591
Db	169447	CCG3AAAACCCCGCGCATTTGAGTAGTATTCATGCTTGAATACAGGCTCGGCGCGAC	169505
QY	592	AATGAATTTGAATCATCGCACCGAAGGATCTCGACTCGCTGATCCCAAGGAAGA	651
Db	169506	AATGAATTT- AACTGATCGCACCGAAGGATCTCGACTCGCTGATCCCAAGGAAG-	169563
QY	652	TTTGTGACTAATGATCCGATGACGCTGCTGCTGCTGACGCGCGGTGCGCGGATCGCC	711
Db	169564	TTTGTGACTAATGATCCGATGACGCTGCTGCTGCTGACGCGCGGTGCGCGGATCGCC	169622
QY	712	TACGTGCGCTCATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCTGCTG	771
Db	169623	TACGTGCGCTCATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCTGCTG	169681
QY	772	CGCGCTTACGATGATCCGCGCGGTTTATGCTTATATACGAAAGATAAGCTG	831
Db	169682	CGCGCTTACGATGATCCGCGCGGTTTATGCTTATATACGAAAGATAAGCTG	169740
QY	832	CGCTGAAGTACAGCTGCTGATCACTCGCTGACGCTGATTTTGTGAGTGGTAAA	891
Db	169741	CGCTGAAGTACAGCTGCTGATCACTCGCTGACGCTGATTTTGTGAGTGGTAAA	169799
QY	892	TTGTTTACAGATGACGCGCGCGGAAAGAGATAA	930
Db	169800	TTGTTTCA-GAGATGACGCGCGCGGAAAGAGATAA	169837
RESULT 9			
LOCUS	AL627278	258050 bp	DNA linear BCT 04-JUL-2003
DEFINITION	Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18, complete chromosome, segment 14/20.		
ACCESSION	AL627278	AL513382	
VERSION	AL627278.1	GI:16504263	
KEYWORDS			
SOURCE	Salmonella enterica subsp. enterica serovar Typhi		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.		
REFERENCE	1 (bases 1 to 258050)		
AUTHORS	Parhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G., Sebahia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connor, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagals, K., Krogh, A., Larsen, T.S., Leather, S., Moulé, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.		
TITLE	Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18		
JOURNAL	Nature 413 (6858), 848-852 (2001)		
MEDLINE	21534947		
PUBMED	11677608		
REFERENCE	2 (bases 1 to 258050)		
AUTHORS	Parhill, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-OCT-2001) Submitted on behalf of the Salmonella sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK		
COMMENT	E-mail: parkhill@sanger.ac.uk		
	Notes: Details of S. typhi sequencing at the Sanger Centre are available on the World Wide Web.		

FEATURES	source	(URL, http://www.sanger.ac.uk/projects/s_typhi/).
	Location/Qualifiers	
	1..258050	/organism=Salmonella enterica subsp. enterica serovar Typhi
	/mol_type="genomic DNA"	
	/strain="CT18"	
	/db_xref="taxon:90370"	
gene	69..1544	
	/gene="tolC"	
	/note="synonym: STY33364"	
CDS	69..1544	
	/gene="tolC"	
	/note="Similar to Salmonella enteritidis outer membrane protein TolC precursor TolC SALEN (Q54001) (491 aa) fasta scores: E(): 0.99.0% id in 491 aa, and to Escherichia coli outer membrane protein tolC precursor tolC or mcb or muka or refi SW:TOLC_ECOLI (P02930) (495 aa) fasta scores: E(): 0.89.7% id in 495 aa Orthologue of E. coli tolC (TOLC_ECOLI); Fasta hit to TOLC_ECOLI (495 aa), 90% identity in 495 aa overlap"	
	/codon_start=1	
	/transl_table=11	
	/product="outer membrane protein TolC precursor"	
	/protein_id="CAD07712.1"	
	/db_xref="GI:16504264"	
	/db_xref="GOA:Q8Z3N8"	
	/db_xref="TrEMBL:Q8Z3N8"	
	/translation="MQMKLLPILGLSGFSTLSQAEINLMQVYQOARLSNPELRKS ADRAAPFKINEARSPILLPOLGLADVTYSNGYDANGINSETSASJLTOTLFDMA SKWRGLTQEKAGIQDVTQDQTLTANTAYKVLNAIDVLSYTAQOEALYRQ LDOITRFNVGVAITDVONARAQDVTLANEVTARNLDNAEISLROVTGNYVELA SLNVEHFKDKPKNALDKAENRNLSLQARLSQDLAREIQIRQDGLHPLTNLTA STGIDTSYSGKSTNSAQYDDSMQNGIKGLNLSPLFYGGMVNSQVQAOQNFVGAS EOLSAHRSVQVTVRSFNINASISINAYKQAVYSAOSSLDVAEAGYSVGTRTIVD VLDATTTLYDAKQQLANARYTILNOLNIKVALTGNEODLLALNLSTLKGPIPTSPES VAPETPEODAAADGYNHSAAPAVQPTAARANSNNGNPRH"	
gene	1757..2428	
	/gene="STY3365"	
CDS	1757..2428	
	/gene="STY3365"	
	/note="Orthologue of E. coli ygiB (YGI_B_ECOLI); Fasta hit to YGI_B_ECOLI (234 aa), 97% identity in 223 aa overlap. Contains a possible N-terminal signal sequence."	
	/codon_start=1	
	/transl_table=11	
	/product="possible lipoprotein"	
	/protein_id="CAD07713.1"	
	/db_xref="GI:16504265"	
	/db_xref="TrEMBL:Q8XG21"	
	/translation="VKRTKSIHHSFRKSWSAHLPVALAVTAVFVLACSEKSDSTV SLYQADDCSAANPGKASCTTAYNNALKEARTAPKATREDCAEFCGCGQOAPA QGMAENQAQACSCSGSFWFMIMAGYMGRLMGCGAGFAQAPLSSKNPSPAYGKY TDAAGKYGAAQPGRTMTVPKTMAPKPAITIVTIVGGFSGVAKQSTVQRSAAAGTST RSMGG"	
misc_feature	1835..1867	
	/gene="STY3365"	
	/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"	
gene	2434..3397	
	/gene="STY3366"	
CDS	2434..3397	
	/gene="STY3366"	
	/note="Fasta hit to YJFC_ECOLI (387 aa), 50% identity in 392 aa overlap Orthologue of E. coli ygiC (YGI_C_ECOLI); Fasta hit to YGI_C_ECOLI (386 aa), 94% identity in 386 aa overlap"	
	/codon_start=1	
	/transl_table=11	
	/product="conserved hypothetical protein"	
	/protein_id="CAD07714.1"	
	/db_xref="GI:16504266"	
	/db_xref="TrEMBL:Q8XFP6"	

```

/translation="VERVSI TERPWRDKATEYGFNFTMYGEPYWCEDAYKLTIAQ
VEKLSDVTAELQMCILKVVVERVIADELMTRIPKHTWGFVRSWQOQQPSLSYRLD
LAWDGIQPKLENNADPTSLYEAATQWLEQDQINAGNLPESDQFSLQELKLE
RFAELRQYFOLLHLTCRDTVEDRGTQIQDCAAEAEIATEFLYIDDLGLGKGG
FIDLQOQIANFLKLYPWFMLREMGSTKLEDAQVRLPEPAWKSIISNALLPLIEM
PDHNLPAFAEDBEHDMKVVKPIPSREGANVSIENKGTIESVEGYPGEGMI
VQOFPPLKFGDSYTLIGSLINDQFAGIGREDRALITQLDSRFYPIHFVEG"
complement(3659..4489)
/gene="STY3367"
/notes="STY3367"
complement(3659..4489)
/gene="STY3367"
/notes="STY3367"
to YGID_ECOLI (271 aa), 87% identity in 266 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07715.1"
/db_xref="GI:16504267"
/db_xref="GOA:Q823N7"
/db_xref="TrEMBL:O823N7"
/translation="MVMLTKLSCXNDIMSLTCLPFLGHGSPMVLDDNDYTRWR
RIGELAPRQALIVVSARWYINGTGTWATERPQTLHDFGGFQALYDHYPAQSPAL
AORVELLAPVVALDKAWGPDHSGWGLIKMYENADIPMVQLSVDSVTKPAWHFEM
GRKLATLDRGVMVASGNVHNLRTRVHGNDIPYWAASFNDVFKANLTWQGFVEQ
HPLVNYLQHEGGLSNPTPEHFLPLLYVLGAWDGKEPITIPVDGIEMSGISMLSVQVG"
4588..5361
/gene="STY3368"
4588..5361
/gene="STY3368"
/notes="Orthologue of E. coli ygie (YGIE_ECOLI); FastA hit
to YGIE_ECOLI (257 aa), 93% identity in 257 aa overlap.
Contains multiple possible membrane spanning hydrophobic
domains"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD07716.1"
/db_xref="GI:16504268"
/db_xref="GOA:Q8XGR4"
/db_xref="TrEMBL:O8XGR4"
/translation="MSVPLLIITLLAGATFTGAFGLVGLQKPSNRVLAFSLGPAAGIM
LILSIEMLPALADTGMSPVLGYNFIIGLGYGLGLRLLPHAPQDLVQKROQPLP
GSIKTALLITGLSHNPEGIATFTVASSNLELFGIALVALHNLPEGLVAGVP
YAATGSKRTATFWAGISGVAELIGVLAWLILGLSVSPVMAAIAVAGINVALSV
ELMPLAKHIDPNNPSYGVLCGMSIMGLSLVILQIGIG"
6004..7800
/gene="STY3370"
6004..7800
/notes="Synonym: asst"
/gene="STY3370"
/notes="STY3370"
/notes="Similar to Klebsiella pneumoniae arylsulfate
sulfotransferase asst TR:P97036 (EMBL:U32616) (598 aa)
fasta scores: E(): 0, 86.8% id in 598 aa"
/codon_start=1
/transl_table=11
/product="probable arylsulfate sulfotransferase"
/protein_id="CAD07717.1"
/db_xref="GI:16504269"
/db_xref="GOA:Q823N6"
/db_xref="TrEMBL:O823N6"
/translation="WEDQYKTLILAGAVLTCGLTAASTPAAGFQPAQPAQKLGAVV
DIGNAPITVLELDSHISIDVVKVTHGKRGVPTTVVGKESLETVDGIPFGLYQ
KFANVTVEYKNGAMKDDYVQTSIAVNHVMDRSISDLOQTKVIVKVPFEDRLY
LYNTFTFPQAEFPHWGEKDKNAGILDAGPAGALPFDIAPYTFVVDTOGEYRWLD
QDTFVNGHDMINKRGYLMGIRTEPRGFTTAVQGOHWYEFDMGQILADHKLPRGFLD
ASHESITVNGVLVRGKRDYRKEDGIVHVTIDQITTEVDKSGRVVDVWDLTKILD
MEDALLDAGACVNDLAHAGQAKLEPDTPYGDALGVGAGRNWAVNSIADAK
DSDILSRHQIVIGIKRDQVFWILAPSKGNKQLASKLLKLPVDDHDKPLTCBENGK
KQDQDFDFTYTHQWLSKGLTVDNGDGRGLQPALPNKSRFVEYKIDFKNGY
VQOWEYKGERGYDYSITSVVYQKORDTRMFGFGSINLFDVGKPTVGLKNEIDYK
TKEVKVEIDLSDKRNQTHYRALLVHPITQMFK"
7820..8491

```

gene

CDS

gene

CDS

gene

CDS

gene

```

/translation="dsba"
/notes="synonym: STY3371"
7820..8491
/notes="dsba"
/notes="Similar to Enterobacter amnigenus disulfide
isomerase dsba TR:09XDPI (EMBL:AF012826) (222 aa) fasta
scores: E(): 0, 90.1% id in 222 aa"
/codon_start=1
/transl_table=11
/product="disulfide isomerase"
/protein_id="CAD07718.1"
/db_xref="GI:16504270"
/db_xref="GOA:O823N5"
/db_xref="TrEMBL:O823N5"
/translation="MSKWTISLFSKSVLITAAVTPFAASAFTEGTDYMWLEKPIPA
DKTLIKVSIACPFCKYDKAVTPVSDKVDLVTFTPEHLETKGYSQASVFAVL
IAKDKAAGISLFDKASKYKFAKFAWYTHYDKERWSDGKDPAAFIKTGLDAAGMSQA
DPEAALKDPAVQSTLEKRWAAVYAKIQGVPAVYVNGKYLITFKIKSIDSMALVRE
LATKK"
misc_feature 7847..8488
Query Match 72.1%; Score 670.8; DB 1; Length 258050;
Best Local Similarity 82.6%; Pred. No. 5.5e-163;
Matches 768; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
QY 1 ATGGAACGACATAAAACGATGTCGGTGTTCGCCAAGTAGTGTGAATTTGGCTCTTTTACC 60
Db 180896 ATGGAACGATATAAACGAAATGTCGGTATTCGCCAAGTGTTCGAGTTTGGCTCTTTTACC 180955
QY 61 GCGCCGCCAGACGAGCTACAGATGAGCGTTTCCTCATCATAGTCAGACGGTATCAAAACTG 120
Db 180956 GCGCAGCCAGACGAGCTGCAATGAGCGTATCTCCATCAGCAGACTGTGGGAAACTG 181015
QY 121 GAGATGAGTTCGAGGTAAAGCTGTTAAACCGTAGCACGACGAGCATTTGSCCTGACCGAA 180
Db 181016 GAGATGAGTTCGAGTTAAATTTGCTACCGCAGCAGCAGCATTTGGGCTCACCGAA 181075
QY 181 GCGGTGAGATTTACTACAGGCTGCGGTGATGCTTCATGAAGTGCAGGATGTTTCAT 240
Db 181076 GCTGGAATAATCTATTATCAGGCTGTCGTGATGCTGATGAAGTGCAGGATGTTTCAC 181135
QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGACGCTACGCATTGCTGCTTCTTCA 300
Db 181136 GAGCAGCTTTACGCTTTTAAATACAGCCTATCGGTACATTCGGTATCGGCTGCTTCTCA 181195
QY 301 ACTATGACACAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360
Db 181196 ACCATGCGGAAAATGCTGCGCGGGCTTACCGCAAAATATTATAAGAGTATCCCGGT 181255
QY 361 TTGAGCGTCAATCTGTTTACCGAAATTCAGCCCCCGACCTGATTCGCCGACGGTCTGGAT 420
Db 181256 CTGCGGCTCAATCTGTTGACAGAAATTCGCGCCCGGATTTAATTCGCCACGGCTGGAC 181315
QY 421 GTGTGATCCGCTCGCGCGGTTGCAAGGATTCAGGCTGTTTCCCGCGCTCTGGGCGG 480
Db 181316 GTGTGATCCGCTGCGCGCGCTGCAAGACTCCAGCTGTTTCCCGTGCAGCTCGGCGG 181375
QY 481 ATGCCAATGTTGTTGTCGCGCGGAAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
Db 181376 ATGCCGATGTTGTTGTCGCGCGGCAAGCCCTACCTCGTCAGTACGGCGTCCCGGAAAAA 181435
QY 541 CCGCGCGATTTGAGTAGTCAATTCATGGCTGAATACAGCGTGGCGCGCCGACRATGATTT 600
Db 181436 CCGCGCGATCTCAGTAGCCCAATTCCTGGCTGGAATATAGCGTGGCGCGCGATATAGTTT 181495
QY 601 GAATGATCGCACCGGAAGGATCTCGACTCGGCTGATCCCAAGAAAGATTTGTGACT 660
Db 181496 GAATTTATGGCGCGGAAGAAATTTCTACCGACTGATTTCCACAGGGGCGGTTTCGTCAT 181555
QY 661 AATGATCCGATGAGCTGTGCGTGCCTGAGCGGGGTGCCGGGATCGCTACGTCGCG 720
Db 181556 AACGATCCGATGAGCTGTGTCGCTGACGTCGCTGACCGCGCGGAAACCGGAAATCGCTTATGTCGCG 181615

```

CDS

QY 721 CTGATGCGGTGATCAACAGAGATCAATCGTGGGAGCTGGAGATCCTGTCGCCGCTTAC 780
 Db 181616 CTGATGCGGTGATCGATGAGATCAATCGTGGGATCGGATCTGCTGCTGCTAT 181675
 QY 781 CAGTCAGATCCAGCCCGGTTTATGGCTTATATACCGAAAGATAGTCCGCTGAAG 840
 Db 181676 CAGTCAGATCCCGCTGCGGTGATCGCGCTGTACCGGAAAGATAGTCCGCTTAA 181735
 QY 841 GTACAGCTGATCAATCACTCGCTGACCGGATTATTGTTGAGCTCGGTAATTTGTTCCAG 900
 Db 181736 GTCCAGTAGTATTAACCGCTTAAGTCTGTTGCTGATGTCGACACCTGTTCCAG 181795
 QY 901 GAGATCAGCGGCGCGGAAAGAGATTA 930
 Db 181796 GGGATCAGCGGCGGAGGAAAGAAATTA 181825

RESULT 10
 AE016845
 LOCUS
 DEFINITION Salmonella enterica subsp. enterica serovar Typhi Ty2, section 12
 of 16 of the complete genome.
 ACCESSION AE016845
 VERSION AE016845.1
 KEYWORDS GI:29139182
 SOURCE Salmonella enterica subsp. enterica serovar Typhi Ty2
 ORGANISM Salmonella enterica subsp. enterica serovar Typhi Ty2
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Salmonella.
 REFERENCE 1 (bases 1 to 299991)
 AUTHORS Deng, W., Liou, S.-R., Plunkett III, G., Mayhew, G. F., Rose, D. J.,
 Burland, V., Kodoyianni, V., Schwartz, D. C. and Blattner, F. R.
 TITLE Comparative Genomics of Salmonella enterica Serovar Typhi Strains
 Ty2 and CT18
 JOURNAL J. Bacteriol. 185 (7), 2330-2337 (2003)
 MEDLINE 22531367
 PUBMED 12644504
 REFERENCE 2 (bases 1 to 299991)
 AUTHORS Deng, W., Liou, S.-R., Plunkett, G. III, Mayhew, G. F., Rose, D. J.,
 Burland, V., Kodoyianni, V., Schwartz, D. C. and Blattner, F. R.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-2002) Laboratory of Genetics, University of
 Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
 FEATURES
 source
 1..299991
 /organism="Salmonella enterica subsp. enterica serovar
 typhi Ty2"
 /mol_type="genomic DNA"
 /strain="Ty2"
 /serovar="Typhi"
 /sub_species="enterica"
 /db_xref="taxon:209261"
 complement(83..709)
 /gene="ftsJ"
 /locus_tag="t3214"
 complement(83..709)
 /gene="ftsJ"
 /locus_tag="t3214"
 /note="corresponds to STV3475 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="cell division protein"
 /protein_id="AAC07050.1"
 /db_xref="GI:29139183"
 /translation="MTGKRSASSSSWLQEHFSDKYVOAKKGLRSPAFKLEIQ
 SKLKPGMTVDLGAAPGWSQYVVTQIGKGRLLIACDLPMDFIVGVDFLQDFRD
 ELVMKALDERVDSKVQVWMSDAPNMSGTPAVDIPRAVYLVEALDMCRDVLAPGGS
 FVVKVFGDFEYLRISLFTKVKVRKPPSSRRARSREVIVATGR"
 797..1129
 /locus_tag="t3215"
 797..1129
 /locus_tag="t3215"

gene
 CDS
 /note="corresponds to STV3476 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAO70751.1"
 /db_xref="GI:29139184"
 /translation="MTRFOSQRQKYTNLSTKQKQHLKGLAHPKPVVLMGNGLTE
 GVLAETQEAHELHKLKVIASEDRKTLIVDAIVRETGACNVQVIGKTLVLRPTK
 ERKISLPR"
 complement(1285..1761)
 /gene="greA"
 /locus_tag="t3216"
 complement(1285..1761)
 /note="corresponds to STV3477 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="transcription elongation factor"
 /protein_id="AAO70752.1"
 /db_xref="GI:29139185"
 /translation="MQAIPMTLRGAKREELDFKSVRRPPIIAIAEAREHGDLKE
 NAEYHAARQQGFCGRIKDIEAKLSNAQVIVDTKMPNNGRVIFGATVTLNLDTEE
 QYTRIVGDDDEAUFKQNLISVNSFIARGLIGKEQDDVVVKTGGDVEVLEKVEYL"
 2009..3442
 /locus_tag="t3217"
 2009..3442
 /note="corresponds to STV3479 from Accession AL513382:
 Salmonella typhi CT18; D-alanyl-D-alanine
 carboxypeptidase"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein"
 /protein_id="AAO70753.1"
 /db_xref="GI:29139186"
 /translation="MRFSEFIILGTSTIAFSVQAANIDEVIKOLPAGANLALMVOKIG
 APAPDIHSQMALPASQKVITALLALILQLGPFRTTLETGKGVNDGLKGDLI
 ARFGDPTLRQDIRMVATLKSGVTQIDGNLIDTISIFASHDKAPGPNNDLTCQF
 SAPAAAIYDRNCFVSLSAQKPNDAFIRVASYPVTMFQSVTLPRGSADAYQE
 LDVPGDLNRYTIGCLPQADPLPLAFIQDGSYAGAILKQELKEAGITVYRGLLR
 OTQVNEPGTIVASKQSAPLHLLKIMLKSDNMIADTVPRMIGHRVFNVPWTRAGSD
 AVQILRQQAQIDIGNTIADGSLSRHNLIAPTMMQVLYIAOHDNELNFISSLPL
 AGYDGLQYRAGLHQGVGKYSAGTSQGVNLAGFITTASGQEMAFVQVLYSGYAV
 PPADQNRRIPLVRFESRDIYQNN"
 complement(3573..4745)
 /locus_tag="t3218"
 complement(3573..4745)
 /note="corresponds to STV3480 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="probable GTP-binding protein"
 /protein_id="AAO70754.1"
 /db_xref="GI:29139187"
 /translation="MKFVDEASILVAGDGGNGCVSFRREKVIKPGDGGDGGGG
 VWWEADENLTLIDYRFEKSPRAESGONGASDCTGKRGKDVTKVPGVTRVIDOGTG
 ETNGDMTKHGRLLVAKGWHGLGNTRFKSVNRPRTQKNTGPGKRDLLLEMLLA
 DVMGLMPNAGSTFIRVSAAPKPVADYFFITLVPSLVGVRMSKSVFVADIFGLI
 EGAAEGAGLIRFHLERLVLLDIDIDPSDPVENARIIIGELEKYSQDLAAK
 PRWLVPNKIDLMKSEAEKAKAIALGWEKGYLISAAQSLQVLDLCKDVMYTFIE
 NPTAQAEAKQKVEFMWDDVHRQQLAEVEEDADDDDDDDDEDEGEVFIYKR"
 complement(4761..5726)
 /locus_tag="t3219"
 complement(4761..5726)
 /note="corresponds to STV3481 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11

/product="putative membrane protein"
/protein_id="AAO70755.1"
/db_xref="GI:29139188"
/translation="MKQAGIGILLATTCMCWGPALPIAMKQVLEWEPSTIVYRFL
MASIGLGLAVLRKLPPIRFRKRWLVLLAATCGLFGNFIFFSSLSQLSPASQ
VIGQSPVGMVAVSFILKBRKQGVIGALMLSLGVMFFNTSLIEFTLRLTYWG
VIFGVGAAMVSYGAKVLRRLASQIILFLYTLCTIALLLPLAKPMVIAQLSDWQ
LACLFPGCLNTLVGCAAEARWAAQVSAIITLPLTLLFDLILSMWPPFFAR
PVLNLLYLGAFFVAVAGMYSAGHRIWGGLRKHETVVSQPRSGE"
complement(5856..6113)
/gene="rplM"
/locus_tag="t3220"
complement(5856..6113)
/gene="rplM"
/locus_tag="t3220"
/notes="corresponds to STY3482 from Accession AL513382;
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L27"
/protein_id="AAO70756.1"
/db_xref="GI:29139189"
/translation="WAHKKAGSTENGSDSEAKRLGVKRGGEAVLAGSIIVRQGTG
FRAGNVGGRDHTLPAKADGKVEVGPKNRKYISVAE"
complement(6133..6444)
/gene="rplU"
/locus_tag="t3221"
complement(6133..6444)
/gene="rplU"
/locus_tag="t3221"
/notes="corresponds to STY3483 from Accession AL513382;
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L21"
/protein_id="AAO70757.1"
/db_xref="GI:29139190"
/translation="WFAVPSGKGKRVSEGTVLEKLDIATGTIFPAEVLMIANG
EEVKGVPVDSGVIAEVAHGRKEKIVFRRRKHRRKQOGRHQPDTDKITGIS
A"
6703..7674
/gene="ispB"
/locus_tag="t3222"
6703..7674
/gene="ispB"
/locus_tag="t3222"
/notes="corresponds to STY3484 from Accession AL513382;
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="octaprenyl-diphosphate synthase"
/protein_id="AAO70758.1"
/db_xref="GI:29139191"
/translation="WNLEKINELTAQDMAGVNATILEQLNSDVQLINQLGYIISGGG
KIRPMIAVLAARAVYQGNHVTIAALTEFHTATLLHDDVDVDSNMRGKATANA
FGNAASVLGDFITRAFOFMTSLGSLKYLEVSAVNVIABEVLQMLNVPDITE
ENMYRVIYKTRLEPAAQCSGILAGCTPEQEKGLQDXYGLYGFALQIOLLDLYSA
DEHLKNNVGDNLNEGKPTLPULLHMRHGTPEQSAMIRTAIEQGNRHLLEPVLAMT
TCGSUWTKQRAEEBAKSAISALQILPDTFWREALIGUAHIAVQDR"
7907..8194
/gene="nlp"
/locus_tag="t3223"
7907..8194
/gene="nlp"
/locus_tag="t3223"
/notes="corresponds to STY3485 from Accession AL513382;
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="Ner-like regulatory protein"
/protein_id="AAO70759.1"
/db_xref="GI:29139192"
/translation="MESKLIIDWHPADIAGLRKKGTSMASRRNGLSSSTIANALTR

Query Match		72.1%; Score 670.8; DB 1; Length 299991;
Best Local Similarity		82.6%; Pred. No. 5.6e-163;
Matches		768; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy	1	ATGGAACGATATAAACCCATGTCGGTGTTCGCAAAAGTAGTGAATTTGGCTCTTTTACC 60
Db	65727	ATGGAACGATATAAACCCATGTCGGTGTTCGCAAAAGTAGTGAATTTGGCTCTTTTACC 65786
Qy	61	GCCGCCCCACAGAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGGTATCAAAACTG 120
Db	65787	GCCGCCCCACAGAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGGTATCAAAACTG 65846
Qy	121	GAAGATGAGTTGCGAGGTAAAGCTGTTAAACCGTAGCACACGCGACATTGGCCCTGACCGAA 180
Db	65847	GAAGATGAGTTGCGAGGTAAAGCTGTTAAACCGTAGCACACGCGACATTGGCCCTGACCGAA 65906
Qy	181	CCCGGTAGAATTACTACAGAGGCTGCCGTGCTATGCTTTCATGAAGTCAGAGTAGTTTCAT 240
Db	65907	GCTCGAAAATCTATTATCAGGCGTGTGCTGCTATGCTGATGAAGTCAGAGTAGTTTCAC 65966
Qy	241	GAGCAACTGTATGCTTCAATACACCCCATCGGACGCTAGGCATTTGCTGTCTTCA 300
Db	65967	GAGCAGCTTTACGCTTTTAAATACACGCTATCGGTACATTCGTATCGGCTGTCTTCA 66026
Qy	301	ACTATGGCACAAAATGTTCTCGCGGGCTGACGCCAAAATGCTGAAAAGATAACCCAGGT 360
Db	66027	ACCATGGCGAAAATGTTCTCGCGGGCTTACCGCAAAAATATTAAAGAGTAGTATCCCGGT 66086
Qy	361	TTGAGCGTCAATCTGTTTACCGGAAATTCAGACGCCCGACCTGATTCGCGACGGTCTCGAT 420
Db	66087	CTGCGGTCATCTGTTGACAGAAATTCGCGCCCGGATTAATTTGCGACGGGCTGGAC 66146
Qy	421	GTGTGATCCGCGTCGCGCGTTCAGGATTCCAGACTGTTTCCGCCCTCTCTGGGCGCG 480
Db	66147	GTGTGATCCGCGTCGCGCGTTCAGGATTCCAGACTGTTTCCGCCCTCTCTGGGCGCG 66206
Qy	481	ATGCCAATGTTGTTGCGCGCGGAAAGCTATCTACAAATACGCGCATACCGGAAAAA 540
Db	66207	ATGCCAATGTTGTTGCGCGCGGAAAGCTATCTACAAATACGCGCATACCGGAAAAA 66266
Qy	541	CCCCCGGATTTGAGTAGTCATTTCATGGCTTGAATACAGCGTGGCGGCCGCAATGAATTT 600
Db	66267	CCCCCGGATCTCAGTAGCCATTCTCGCTGGAATATAGCGTGGCGCCGGATATAGATTT 66326
Qy	601	GAATGATCGCACCGGAAGGATCTCGACTGCGCTGATCCACAGAGAAATTTGTGACT 660
Db	66327	GAATGATCGCACCGGAAGGATTTCTACCGGATGATTTCCACAGAGGCGGTTCTGTCAT 66386
Qy	661	AATGATCCGATGACGCTGGTGGCTGAGCGCGGTGCCGGGATCGCTACGTGCCG 720
Db	66387	AACGATCCGATGACGCTGGTGGCTGAGCGCGGATCGCTACGTATGTCGCG 66446
Qy	721	CTGATGTGGGTGATCAACGAGATCAATCGTGGGAGGTGAGATTCCTGTGCGCGCTTAC 780
Db	66447	CTGATGTGGGTGATCGATGAGATCAATCGTGGGATCTGAGATTCCTGTGCGCTTAC 66506
Qy	781	CAGTCAGATCCACCCCGGTTTATCGCTTATATACCGAAAAGATAGCTGCGCTCAAG 840
Db	66507	CAGTCAGATCCACCCCGGTTTATCGCTTATATACCGAAAAGATAGCTGCGCTTAA 66566
Qy	841	GTACAGGTCTGTATCAAACTCGCTCAGCGATTATTTTGTGAGGTGCGTAAATTTTTCAG 900
Db	66567	GTGAGGTGATGATTACGCGTTAACTGACTATTTTGTGATGTCGACACCTGTTTCAG 66626
Qy	901	GAGATGACCGGCGCGGAAAGAGATAA 930
Db	66627	GAGATGACCGGCGGAAAGAGATAA 66656
RESULT 11		
AE008855		
LOCUS		AE008855 21252 bp DNA linear BCT 23-APR-2003

DEFINITION Salmonella typhimurium LT2, section 159 of 220 of the complete genome.

ACCESSION AE008855 AE006468

VERSION AE008855.1 GI:16421920

KEYWORDS

SOURCE

ORGANISM

Salmonella typhimurium LT2

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE

AUTHORS

McClalland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, P., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, B., Sun, H., Flores, L., Miller, W., Stonking, T., Nhan, M., Waterston, R., and Wilson, R.K.

TITLE

Complete genome sequence of Salmonella enterica serovar Typhimurium LT2

JOURNAL

MEDLINE

21534948

PUBMED

11677609

REFERENCE

2 (bases 1 to 21252)

AUTHORS

The Salmonella typhimurium Genome Sequencing Project

CONSTRM

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA

COMMENT

Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs; GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and RegionDB: http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one mi3 subclone.

FEATURES

source

1..21252

/organism="Salmonella typhimurium LT2"

/mol_type="genomic DNA"

/strain="LT2: SGSC 1412; ATCC 700720"

/db_xref="ATCC:700720"

/db_xref="taxon:99287"

/note="LT2"

207..482

/gene="yhcn"

/note="synonym: STM3361"

207..212

/gene="yhcn"

/note="putative RBS for yhcN; RegulonDB:STMSIH003374"

219..482

/gene="yhcn"

/note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 75% identity in aa 18 - 104"

/codon_start=1

/transl_table=11

/product="putative outer membrane protein"

/protein_id="AAL22230.1"

/db_xref="GI:16421921"

/translation="MKIKTTVATPLSLSVLFGAFAPAEPIFAEQQNREAGSVSVSA IGSPMDMNAULSKKADQEQATYHITEARSGSNWHTAELYK"

579..852

/gene="STM3362"

579..584

/gene="STM3362"

/note="putative RBS for STM3362; RegulonDB:STMSIH003375"

586..852

/gene="STM3362"

/note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 53% identity in aa 18 - 104"

/codon_start=1

/transl_table=11

/product="putative periplasmic protein"

/protein_id="AAL22231.1"

/db_xref="GI:16421922"

/translation="MKTIVIIASLGLATLISFGANAAVHQAQNLQPMGTISVS QIGSTPMDMRQEIYAKAEKAGANSYRIIEIKGDNWHTAELYK"

complement(912..1197)

/gene="yhco"

/note="synonym: STM3363"

complement(912..1184)

/gene="yhco"

/note="similar to E. coli orf, hypothetical protein (AAC76271.1); Blastp hit to AAC76271.1 (90 aa), 76% identity in aa 1 - 90"

/codon_start=1

/transl_table=11

/product="putative cytoplasmic protein"

/protein_id="AAL22232.1"

/db_xref="GI:16421923"

/translation="MNVVTFDFNDIKQSDFYREPTOTFGLASEKVDLTLWDAMS DILPLLEIEFVHLDPKLRRYGALILLFDEAEELGRLENVRH"

complement(1192..1197)

/gene="yhco"

/note="putative RBS for yhcO; RegulonDB:STMSIH003376"

complement(1356..3323)

/gene="yhcp"

/note="synonym: STM3364"

complement(1356..3323)

/gene="yhcp"

/note="similar to E. coli orf, hypothetical protein (AAC76272.1); Blastp hit to AAC76272.1 (655 aa), 92% identity in aa 1 - 655"

/codon_start=1

/transl_table=11

/product="putative inner membrane protein"

/protein_id="AAL22233.1"

/db_xref="GI:16421924"

/translation="MGIFSIANQHIRFAVKACAIVALFIFGHFQLEPRWAVLTAA IVAAGPAPAGGEPYSGAIRYGRMLRIIGTFIGCAALIIISIRAPMLILVCCW AGCTWISLVRIENSYAGLSYTAIIIVITOTEPPLTPQALERCSEIVIGCA ILADLLFSRSLKQVDRELDLSVAQVLMOLCKHGDSEVDNAGDVRRTALE GRSNLSNMRKWRFRANRLKALNTLSLTIQSCETYLIIQNTRELIIDTREFLEF PVEVDQVHRQRMRIVITGERTPVTLISWGAATRYLLLRGVISNTKISATE EEILOGEPVVKVESAEERHAMVFNWRTLLSCILGTLFWLWTGNGAMVIAVVS LAMRPNRMVCDITGYTLAALPLGLLYFLVITPNTQSMLLCLSLAVLFFGIE VOKRGLSGMALSTINIIIVLDNPMTHFSQFLDSALGOIVGCMIAFVILLVRDPSK DRTGRVLLNQFVSAASVAMTNVVRKENRLPALYQQLFLMNKPGDLPKRFALTMT IIAHOSLRDAPIVNEDLSVFHRLQERTADHVISAGSDDKRRRYFQLLDLIDYQEK LRIWEAPQVTFVRKLTGNLHKYQVALTDS"

complement(3329..4270)

/gene="yhco"

/note="synonym: STM3365"

complement(3329..4261)

/gene="yhco"

/note="similar to E. coli putative membrane protein (AAC76273.1); Blastp hit to AAC76273.1 (310 aa), 92% identity in aa 1 - 310"

/codon_start=1

/transl_table=11

		/product="putative membrane located multidrug resistance protein"
		/protein_id="AAU22234.1"
		/db_xref="GI:16421925"
		/translation="MKTLTKLSRTAITLVILAFIAFRAMWYVYTESPWTDRARFS ADVEAGPRNRIQVQMSREIIOQANNVLOTVLHOLAKAQAOTDLAKDLERTVIRAP EKQAEAGRNRIQVQMSREIIOQANNVLOTVLHOLAKAQAOTDLAKDLERTVIRAP ADGWYNLNYAGFEITRGTATVALYKNSFYQAYMEETKLEGVPGFPAEITPLGS NRVKGTVDSVAGVNNASSTDAKGNATDSNLEWVRLAQRVFPVIRLDEQQQGNLWLP AGVATVITGKQDRDASQSFRRKLALHRLREFG"
RBS		/complement(4265..4270) /gene="yhqQ" /note="putative RBS for yhcQ; RegulonDB:STWS1H003377"
	gene	/complement(4269..4481) /gene="yhcr"
		/note="synonym: STM3366"
	CDS	/complement(4269..4472) /gene="yhcr" /note="similar to E. coli orf, hypothetical protein (AAC76274.1); Blastp hit to AAC76274.1 (90 aa), 98% identity in aa 24 - 90"
		/codon_start=1 /translation_table=11 /product="putative inner membrane protein"
		/protein_id="AAU22235.1"
		/db_xref="GI:16421926"
		/translation="MSLPFVIVVGLSFPPIFFELLLSLAFWLVRMLVPTGIYDFV WPALENTALYCCGLFYLISRLFV"
RBS		/complement(4476..4481) /gene="yhcr"
	gene	/note="putative RBS for yhcR; RegulonDB:STWS1H003378"
		4654..5583
		/gene="yhcrS"
		/note="synonym: STM3367"
	CDS	4654..5583 /gene="yhcrS"
		/note="similar to E. coli putative transcriptional regulator LYSR-type (AAC76275.1); Blastp hit to AAC76275.1 (309 aa), 95% identity in aa 1 - 309"
		/codon_start=1 /translation_table=11 /product="putative LysR family transcriptional regulator"
		/protein_id="AAU22236.1"
		/db_xref="GI:16421927"
		/translation="MERLKMSYFAKVVERGSFTAAARQLQMSVSSIQTAKLEDEL QVKLLNRSTISGLTAGKIYYQCCRMHEVDVHEQLYAFNFTIGTLRIGCSSTM AQNVLAKLKEVGLAVNLVTGI PADLJADGLVYRVNGLDGLSLRSLRGLA MPMVCAAKPYLAQGVPEPADLSHSLVSVRPDNEPELLAPGISLRLIPQGRF VTNDPMTLVRLTAGYGIAXVPLMWVIDEINRGDLEILLPRYQSDPFPVALITEKDK LPLKQVIVNALDIYFVDVAHLFGVHGROK"
	gene	/complement(5706..7151) /gene="tldb"
		/note="synonym: STM3369"
		71.68; Score 666; DB 1; Length 21252; Best Local Similarity 82.3%; Pred.No. 8.3e-162; Matches 765; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
Qy	1	ATGGAACGAGCTAAAAACGATGTGGTGTTTGCCAAAGTAGTGTGAATTTGGCTCTTTTACC 60
Db	4654	ATGGAACGATTAAAAAGAAAGTCCGATTTCGCAAGTTCGAGTTTGCTCTTTTACC 4713
Qy	61	CGCCGCCGACAGCTACAGATGAGCGTTTCGTCCTCATCATGTCAGCGGTATCAAACTG 120
Db	4714	CGCCGCCGACAGCTGCAAAATGAGCGGTATCGTCCATCATGACCGACACTGTGGCGAACTG 4773
Qy	121	GAAGATGAGTTGAGGTAAGAGCTGTTAAACCGTAGCACACGACACTTGGCCTGACCGAA 180
Db	4774	GAAGATGAGCTTCAGGTTAAATGCTCAACCGCAGACACGACGATTTGGCTACCGAA 4833
Qy	181	CGCGGTAGATTACTACCGAGGCTCGCGTGTATGCTTATGAAAGTCGAGGATGTTTCAT 240
Db	4834	GCTGGAAAAATCTATTACAGGGTGTGCTGCTATGCTGATGAAGTCGAGGATGTTTCAC 4893

Qy	241	GAGCACTGATCGCTTCAATACACCCCATCGGACCGTACGCATTGCGTCTTCTCA	300
Db	4894	GAGCAGCTTTACCGCTTTTAAATAACACCGCTATCGGTACATTTGCGTATCGGCTGTTCTTCA	4953
Qy	301	ACTATGCGACAAATGTTCTCGCGGGCTGACAGCCAAATATGCTGAAAGATATACCCAGGT	360
Db	4954	ACCATGCGCAAAATGTGCTGGCGGGCTTACCGCAAAATTTATTAAGAGTATCCCGGT	5013
Qy	361	TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCGGACCTGATTCGCGACGCTCTGGAT	420
Db	5014	CTGGCGGTCAATCTGGTGACAGGAATTCGCGGCCCGGATTTAAATGCGGACGGGCTGGAC	5073
Qy	421	GTGGTGATCCGCGCTCGCGGGTTCGACAGGATTCAGAGCTGTTTTCCGCGCGCTCTGGGCGCG	480
Db	5074	GTGGTGATCCGCGTGGGCGCGCTGCAAGACTCCAGCCTGTTTTCCCGTGACTCGGCGCG	5133
Qy	481	ATGCGAATGTTGGTTGTCGCGCGGAAAGCTATCTACACAATACGGCATACCGGAAAAA	540
Db	5134	ATCCGATGTTGTTGTCGCGCGGCAAGCGTACTCGCTCAGTACGGCGTCCCGGAAAAA	5193
Qy	541	CCGCGCGATTTGAGTAGTCAATTCATGGCTTGAATACAGCGTGGCGGCCGACAATGAATTT	600
Db	5194	CCGCGCGATCTCAGTAGCCATTCCTGGCTGGATATAGCGTGGCGCGGATATGAGTTT	5253
Qy	601	GAATGATGTCGACCGGAAGGATCTCGACTCGGCTGATGCCACAGGAAGATTTGTGACT	660
Db	5254	GAATCTATTCGCGCGGAAAGGAAATTTCTACCGGACTGATTCACAGGGCGGTTCTGCAC	5313
Qy	661	AATGATCCGATGACGTGGTGCGCTGGCTGACGCGCGGCTGCGGGATCGGCTACGTGCGG	720
Db	5314	AAGCATCCGATGACGTGGTACGCTGGCTGACCGCGCGGAACCGGATTCGTTATGCGCG	5373
Qy	721	CTGATGTGGGTGATCAACAGAGATCAATCGTGGGAGCTGGAGATCTCTGCTGCCGCGTTAC	780
Db	5374	CTGATGTGGGTGATCGATGAGATCAATCGTGGGAGCTCGAGATCTCTGCTGCCGCTAT	5433
Qy	781	CAGTCAGATCCAGCGCGGTTATCGTTATATACCGAAGAAACATAGCTGCCGCTGAAG	840
Db	5434	CAGTCGATCCCCGTCGCGTGAAGCGCTGTACCGAAGAAAGATAGCTGCCGCTTAA	5493
Qy	841	GTCAGGTGATCAACTCGCTGACGGATTTATTTGTTGAGTCGGTAAATGTTTCAG	900
Db	5494	GTCAGGTAGTGAATTAACGCGTTAACTGACTATTTGTTGATCGCACACCTGTTTCAG	5553
Qy	901	GAGATCAGCGCGCGGGAAGAGAGTAA	930
Db	5554	GGAATGCACGGCGGAGGAAAAAGAAATAA	5583
RESULT 12			
AR389774			
LOCUS	AR389774	933 bp	DNA
DEFINITION	Sequence 6503 from patent US 6610836.	linear	PAT 18-DEC-2003
ACCESSION	AR389774		
VERSION	AR389774.1	GI:40099990	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 933)		
AUTHORS	Bretton, G.L. and Osborne, M.		
TITLE	Nucleic acid amino acid sequences relating to Klebsiella pneumoniae		
JOURNAL	For diagnostics and therapeutics		
FEATURES	Patent: US 6610836-A 6503 26-AUG-2003;		
source	Location/Qualifiers		
	1..933		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	66.8%;	Score 621.4;	DB 6; Length 933;
Best Local Similarity	79.8%;	Pred. No. 2.6e-150;	

Matches	733;	Conservative	0;	Mismatches	186;	Indels	0;	Gaps	0;
Qy	1	ATGGAACGAACTAAACAGCATGTCGGTGTGTTGGCCAAAGTAGTTGAAATTTGGCTCTTTTACC	60						
Db	4	ATGGAACGATTAAACAGCATGTCGGTGTGTTGGCCAAAGTAGTTGAAATTTAGGTTCAITTCACC	63						
Qy	61	GGCCGGCCAGACAGCTACAGATAGCGTTTCGTTCATAGTCAGTACAGAGGGTATCAAAACTG	120						
Db	64	GCAGCGCGCCGCGCAGCTTCAGATAGCGTTTCCTCGATCAGTCAGTACAGACCGTCGCGCAACTG	123						
Qy	121	GAAGATGAGTTCAGGTAAGAGCTGTTAAACCGTAGCACGACAGCATTTGGCCTTGACCGAA	180						
Db	124	GAAGATGAGCTCAGGTAAGAGCTGCTCAACCGAGCACCCGACAGCCTTGGCTTAACTGMAA	183						
Qy	181	GGCGGTAGAAATTTACTACGAGGGCTGGCGTCTGTTATGCTTCAATGAAGTCAGGATGTTCAT	240						
Db	184	GCAGGTAAAAATCTATTACCGGGCTGGCGTAGGATGCTGTTTGAAGCGCAGGATGTCCAT	243						
Qy	241	GAGCAACTGTATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTTGGCTGTGTTCTTCA	300						
Db	244	GAGCAATTATACGCAATTTAAACAATACCCCCATCGGACGCTGGCGCATCGGCTGCTTCCA	303						
Qy	301	ACTATGGCAAAATTTGTTCTGCGGGGTGACGCCAAAATGCTGAAGAATAACCCAGGT	360						
Db	304	ACTATGGCACAGAAATTTGTTGGCGGCATTAACGGCGAAATGCTCAAGAATAACCCGGG	363						
Qy	361	TTGAGCGTCAATCTGGTTACCGGAATTCAGCCGCCGACCTGATTGCCAGCGTCTGGAT	420						
Db	364	CTCTCTGTTTAATCTGGTCACGGCATCTCTGCGCCAGACCTGATGCCAGCGCCTTGAC	423						
Qy	421	GTGGTGATCCGGTCCGGCGGTTGAGAGATTCCAGCCTGTTTCCCGCGCGTCTGGCGCG	480						
Db	424	GTGGTGATCCGGTCCGGCGGCTCGAGGATTCAGACCTTTTTTCCCGCGCGTCTGGGCTCA	483						
Qy	481	ATGCCAAATGGTGGTGTGCGCGCGAAAGCTATCTCACAAATACGGCATACCGGAAAAA	540						
Db	484	ATGCCGATGGTATTATATGCGCGCTATAAAGCTATCTGGCGAGGCTGGCAATCCAGAGAA	543						
Qy	541	CCGCGGATTTGAGTAGTCATTCAATGGCTTGAATACAGCGTGGCGGCCGACCAATGAATTT	600						
Db	544	CCGCGCGACCTTGCCGGAATGCTGCTGGCTGGAAATATAGCTGGCGCGGACCAATGATTT	603						
Qy	601	GAACTGATCGACCGGAGAGGATCTCGACTCGCCTGATCCCAAGAGGAGATTTGTGACT	660						
Db	604	GTCATTATCGCGCCGGAAGGATCTCGACCGCCTGACCCCGCAGGGGCGCTTCGTGACC	663						
Qy	661	AATGATCCGATGACCTGGTGGCTGGCTGACGCGGGTCCCGGGATTCGCTACGTGCGG	720						
Db	664	AACGACCTATGACGCTGGTGGCTGGCTGACCGCGGGGTGGATTCGCTATGTACCC	723						
Qy	721	CTGATGTGGGTGATCAACAGAGATCAATGTGTGGGAGCTGGAGATCTCTGTGCGCGGTTC	780						
Db	724	TTAATGTGGCGATCGAAGAGATCAACCGCGGTGAACTTCAGATCTCTGTGCGCCAGCTAT	783						
Qy	781	CAGTCAGATCCAGCCCGGTTTATGCGTTATATACGAAAGATTAAGCTCGCGCTGAG	840						
Db	784	CAGTCGATCCCGGACCGGTTCTAGCCCTGTATACGAAAGGATTAATCTGCGCGTGAA	843						
Qy	841	GTACAGGTCTGATCAACTCGCTCAGGATTTATTTGTTGAGTCTGGTAAATTTGTTTCAG	900						
Db	844	GTGAGGTCTGATCAACTCTGACGGAGTATTTTGTTCGACGTGGCGAAGATCTATCAG	903						
Qy	901	GAGATGACGCGCGCGGA	919						
Db	904	GGCATGACGCGCGCGGA	922						

RESIST 13

RESULI 13
BX950851 03/c

WPCOMMENT
BX30831

Sequence split into 51 fragments LOCUS BX950851 Accession BX950851

Fragment Name

BX950851_00

Continuation 74 of 51) of BX950851 from base 300001 (BX950851 *Erwinia carotovora* subsp. *carotovora*)

Query Match

Query Match
Best Local Similarity
Best Local Score
Best Local Pred. No.

BEST LOCAL SIMILARITY 70.3%, FREQ: NO. 1:30 131,
 Matches 676: Conservative 0: Mismatches 210: Indels 0:

1 ATGGAACGACTAAACGCGATGTCGGTGTTCGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60

db 25479 ATGGAAAGACTAAAGAGTATGTCGGTGTTCGCCCGCTGGTGGAA TGTGGTTCCTTTACC 25420

61 GCCGCCGCCAGACAGCTACAGATGAGCGCTTCGTCCATCAGTCAGACGGTATCAAAACTG 120

db 25419 GCCGCCGACGCCAGTTGCAGATGAGGGTATCCGCCGTCCGCCAGACCGTCTCCAAGCTT 25360

121 CAAGATGAGTTGCACGTAAGCTGTAAACCGTAGCACAAGCAGCATTTGGCCTGACCGAA 180

25359 G A A G A T C A C A G A T T A G C T G T G A A C C G C A G T A C G G C A G T A T T G G T C T G A C G G A A 25300

181 GCGCGTACGAGTTTACCTACCGCGCTGCCCTCGTATGCTTCA TGAAGTCAGGATGTTCA T 240

25298 GCGGCGAABAAATTTACATACTACCGCTGCCGTAGAAATGCTACATGAGGTCATGAGGTACAC 25240

341 CAGCAGCTGCTCTCAGATATACCCAGCTCCGCTACGCTGGTGTCTTCA 300

```
Db 25239 GAACAGCTTATGCTTCAACACACGCGGATCGGCACGCTACGATTGGTAGTTCGTCT 25180
QY 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGT 360
Db 25179 ACCATGGCGAGATGTGTATCCACCATGACTGCGCGATGCTGAAGNATATCTCGT 25120
QY 361 TTGAGCGTCAATCTGTTACCGGAATCCAGCCCGCGACCTGATTCGCGACGCTTGAT 420
Db 25119 TTGTCCGTCAATCTGTTACCGGCAATTCGCGCTCCGACTGATCGCGACGCTGGAT 25060
QY 421 GTGGTATCCGCGTGGCGGCTGACAGGATTCAGGCTGTTCCGCGCTGTTCCGCGCG 480
Db 25059 ATCGTATCCGCGTGGTGACATACAGGATTCACCGCTGTTCCGCGCTGTTAGGCTG 25000
QY 481 ATGCCAATCGTGTGTCGCCCGCAAGCTATCTCACACAATACCGGATACCGGAAAA 540
Db 24999 ATGCCGATGGTGATGTCGCCGGAAGCTATCTGGCGCAGCAGGACGCGGATAG 24940
QY 541 CCGCGGATTTGAGTAGTCAATTCATGCTTGAATACAGCGGTGCGCGCGACGATTA 600
Db 24939 CAGCGAGATATGTTGAATTTTCTGGCTGGATACAGGCTGCGCGCGACAGTGA 24880
QY 601 GAATGATCGCAGCGGAGGATCTCGACTCGCTGATCCCAAGGAAGATTTCGTGACT 660
Db 24879 GAATGATTCACCGGAGGCTATTCGACCGCGTCAACACACAGGCGCTTGTCA 24820
QY 661 ATGATCCGATGACGCTGTGCGTGTGCTGACGCGGCTGCGGATCGCTACGTCG 720
Db 24819 AACGATCCGCAACGCTGTGCTGCTCAAGATCGGAGCGGCTATCGCTATATCC 24760
QY 721 CTGATGCTGGTGTATCAACAGATCAATCGTGGGAGCTGGAGATCTGTCGCGGTTAC 780
Db 24759 CTCATGTGATTTGGACGAAATCAATCTGGGAAATCGAGATCTTATCAACGCTAC 24700
QY 781 CAGTCAGATCCAGCGCGCTTATCGTGTATATACGAAAAAGATAAGCTGCGCTGA 840
Db 24699 CACTCCGATCCGCGCGGCTACGCGCTTACCGCGCAAGACAACTGCGCGTAAA 24640
QY 841 GTACAGTCTGATCACTCGCTGCGGATTTTGTGTGAGTCTG 886
Db 24639 GTACAGTCTGATTAATATGACGGAATACCTTCAAGAGAGCTG 24594

RESULT 14
ECOTLDD1/c 2788 bp DNA linear BCT 14-APR-2000
LOCUS Escherichia coli DNA for Tidd protein, complete cds.
DEFINITION D4451
ACCESSION D4451.1 Gi:1732436
VERSION Tidd protein.
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (sites)
AUTHORS Murayama,N., Shimizu,H., Takiguchi,S., Baba,Y., Amino,H.,
Horiuchi,T., Sekimizu,K. and Miki,T.
TITLE Evidence for involvement of Escherichia coli genes pmBA, csrA and a
previously unrecognized gene tidd, in the control of DNA gyrase by
tidd (ccdB) of sex factor F
J. Mol. Biol. 256 (3), 483-502 (1996)
MEDLINE 96177756
PUBMED 8604133
REFERENCE 2 (bases 1 to 2788)
AUTHORS Miki,T.
TITLE Direct Submission
JOURNAL Submitted (24-DEC-1994) Takeyoshi Miki, Kyushu University, Faculty
of Pharmaceutical Sciences; 3-1-1, Maidashi, Higashi-ku, Fukuoka,
Fukuoka 812, Japan (E-mail:miki@bisei.phar.kyushu-u.ac.jp,
Tel:092-641-1151(ex.6187), Fax:092-632-5648)
LOCATION/Qualifiers
FEATURES
1. .2788
source
```

```
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12"
/isolate="KP4714"
/db_xref="taxon:562"
/clone="pKP1846"
275. .280
/evidence=not_experimental
297. .618
/notes="91 bp direct repeat; each repeat contains two REP
(repetitive extragenic palindromic) sequences"
/citations=[1]
/rpt_type=tandem
/rpt_unit="345. .435"
298. .303
/evidence=not_experimental
306. .344
/citations=[1]
/rpt_family="REP"
345. .435
/citations=[1]
/rpt_type=inverted
436. .526
/citations=[1]
/rpt_family="REP"
527. .618
/citations=[1]
/rpt_type=inverted
648. .650
/evidence=not_experimental
662. .2107
/genes="tidd"
/function="Modulator of interaction between LetD protein
and DNA gyrase"
/codon_start=1
/transl_table=11
/product="Tidd protein"
/protein_id="BAA07913.1"
/db_xref="GI:1732437"
/translations="MSLNLVSEOLLAKHODLFAILQOLAERLDYGLYFOSSY
HESWLEDRIDKGSNIDQGVRAISGEKTGFAYADQISLALQSAQAARTIVRD
SGDKVQTGAVHSHPLTSDVPLQSMREKDLRLRVKAREADKRVQVETASLS
GEYRLILRAITDGLAADVRLVLSVSVLVEEDKRGKRGAGGGRGFGTFFLADLD
GEYRADAXAEAVRMALVNLASVAPAGTVPVVLGAGPVLHLEAVHGLEDFNRR
GTSVFGQVGLVASELCTVDDGTVDRRGSVAIDDEGTGQYNYLIENGILGYMQ
DKLNARLMGMTPTNGRRRSYAHLPMPMTNTYMLPKGTPOEIIESVYGIYAPNFG
GGVDYITSGKFVFTSEAYLIENGKVTVPKGTATLIGSGIETMQQISWVGNLKLNG
VGVCGRGQSLPVGVQPIKVDNLTVGTA"
2036
/genes="tidd50"
2036
/genes="tidd50"
/citation=[1]
/phenotype="Tolerant to LetD product growth inhibition"
/replace="t"

ORIGIN
Query Match 58.9%; Score 548; DB 1; Length 2788;
Best Local Similarity 100.0%; Pred. No. 3.2e-131;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 383 GAATTCAGCCCCGACCTGATTGCCGAGCTGTGGATGTGGATCCCGCTCGGCGCGT 442
Db 2788 GAATTCAGCCCCGACCTGATTGCCGAGCTGTGGATGTGGATCCCGCTCGGCGCGT 2729
QY 443 TGCAGGATTCAGCCCTGTTTTCGCCGCGCTGCGCGCGATGCCAATGGTGTGCGCCG 502
Db 2728 TGCAGGATTCAGCCCTGTTTTCGCCGCGCTGCGCGCGATGCCAATGGTGTGCGCCG 2669
```

QY 503 CGAAAGCTATCTCACAAATACGGCATACCGGAAAAACCGCCGATTTGAGTAGTCATT 562
 Db 2668 CGAAAGCTATCTCACAAATACGGCATACCGGAAAAACCGCCGATTTGAGTAGTCATT 2609
 QY 563 CATGCTTGAATACAGCGTGGCGCGGACCAATGAATTTGAATCTGATCGCACCGAAGGGA 622
 Db 2608 CATGCTTGAATACAGCGTGGCGCGGACCAATGAATTTGAATCTGATCGCACCGAAGGGA 2549
 QY 623 TCTCGACTCGCTGATCCACAAAGAAATTTGACTAAATGATCCGATGACGCTGTGTC 682
 Db 2548 TCTCGACTCGCTGATCCACAAAGAAATTTGACTAAATGATCCGATGACGCTGTGTC 2489
 QY 683 GCTGCTGACGCGGGTGGCGGATCGCGTACGTGCGCGCTGATGCGGTGATCAACGAGA 742
 Db 2488 GCTGCTGACGCGGGTGGCGGATCGCGTACGTGCGCGCTGATGCGGTGATCAACGAGA 2429
 QY 743 TCAATCGTGGGAGTGGAGATCTCTGCTGCGCGTTCACAGTCAGATCCAGCCCGGTTT 802
 Db 2428 TCAATCGTGGGAGTGGAGATCTCTGCTGCGCGTTCACAGTCAGATCCAGCCCGGTTT 2369
 QY 803 ATGCGTTATATACGAAAGAAATAGTCCGCTGAAGGTACAGGTCTGTGATCAACTGCG 862
 Db 2368 ATGCGTTATATACGAAAGAAATAGTCCGCTGAAGGTACAGGTCTGTGATCAACTGCG 2309
 QY 863 TGACGGATTTTGTGAGTGGGTAAATTTGTTTTCAGGAGATGACGCGCGCGGAAAG 922
 Db 2308 TGACGGATTTTGTGAGTGGGTAAATTTGTTTTCAGGAGATGACGCGCGCGGAAAG 2249
 QY 923 AGAAGTAA 930
 Db 2248 AGAAGTAA 2241

RESULT 15
 AB077386 1512 bp DNA linear BCT 02-FEB-2002
 LOCUS Serratia marcescens tdrA gene for putative temperature-dependent
 DEFINITION regulator A, complete cds.
 ACCESSION AB077386
 VERSION 1
 KEYWORDS Serratia marcescens
 SOURCE Serratia marcescens
 ORGANISM Serratia marcescens
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Serratia.
 REFERENCE 1
 AUTHORS Baba, M., Midorikawa, Y., Nakagawa, Y., Fujita, M. and Matsuyama, T.
 TITLE Serratia marcescens and Escherichia coli genes controlling
 temperature-dependent production of structurally unrelated
 secondary metabolites such as prodigiosin and serrawettin
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1512)
 AUTHORS Nakagawa, Y., Baba, M. and Matsuyama, T.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-2002) Yoji Nakagawa, Niigata University, Faculty
 of Agriculture, 8050, Ikarashi 2-no-cho, Niigata City, Niigata
 950-2181, Japan (E-mail:nakagawa@agr.niigata-u.ac.jp,
 Tel:81-25-262-6360(ex.6360), Fax:81-25-262-6360)
 FEATURES
 source
 1..1512
 /organism="Serratia marcescens"
 /mol_type="genomic DNA"
 /db_xref="taxon:615"
 301..1212
 /gene="tdrA"
 301..1212
 /gene="tdrA"
 /codon_start=1
 /transl_table=1
 /product="putative temperature-dependent regulator A"
 /protein_id="BAB84544.1"
 /db_xref="GI:18478993"
 /translation="MERLRMSVFAKVVFGSFTAAARQLGMSVSSISQTVSKLENEL"

ORIGIN
 Query Match 56.8%; Score 528.2; DB 1; Length 1512;
 Best Local Similarity 74.1%; Pred. No. 4.3e-126;
 Matches 668; Conservative 0; Mismatches 233; Indels 0; Gaps 0;
 QY 1 ATGGAACGACTAAACGCGATGTCGGTGTTCGAAAGTAGTTGAATTTGGCTCTTTTACC 60
 Db 301 ATGGAAGATTAAACGCGATGTCGGTGTTCGAAAGTAGTTGAGTTCGGTCTGTTTACC 360
 QY 61 GCGCGCCGACAGACTACAGATGAGCGTTTCGTTCATCAGTCAGACGGTATCAAACTG 120
 Db 361 GCGCGCCGCGCAACTCGGCATGAGCGTTTCATCGATCAGCAAAACCGTCTCGAACTG 420
 QY 121 GAAGATGAGTTCAGGTAAAGCTGTAAACCGTAGCACACGAGGATTCGCCCTGACCGAA 180
 Db 421 GAAATGAGTTCAGGTAAAGCTGTAAACCGTAGCACACGAGGATTCGCCCTGACCGAG 480
 QY 181 GCGGTGAGATTTACTACACGAGGCTGCCGTCTGTATGCTTCATGAAGTCAGGATTTCA 240
 Db 481 GCGGCAAAATCTACTATACGAGGCTGCCGACGATGCTGCGAAGTACAGCAAGTGCAT 540
 QY 241 GAGCACTGATGCTTCAATACACCCCACTGCGGACGCTACGATTCGGCTGTTCTCA 300
 Db 541 GAAACGCTGATGCTTCAATACACCCCACTGCGGACGCTACGATTCGGCTGTTCTCA 600
 QY 301 ACTATGCGCAAAATGCTTCGCGCGGCTGACAGCAAAATGCTGAAAGATACCCAGT 360
 Db 601 ACCATGCGCAAAACGCTGTCGCGAATGACGCGGAGATGATGAAGAGTACCGGCG 660
 QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGGACCTGATTCGCGACGCTGAT 420
 Db 661 CTGACGCTCAACCTGTTAAACCGGATTCGCGGCGGATCTGATCACCGACGGTGGAT 720
 QY 421 GTGGTATCGCGCTGCGCGGTTGAGGATTCAGGCTGTTTCGCGCGCTGTCGCGCG 480
 Db 721 CTGGTATCCGACAGAGCGCTGAGGACTCAGGCTGTTCTTCGCGCGCTGCGGCGAG 780
 QY 481 ATCCCAATGCTGTGTGCGCGCAAAAGCTATCTCACAAATACGGCATACCGAAAAA 540
 Db 781 ATCCGATGCTGTGTGCGCGCAAAAGCTATCTCACAAATACGGCATACCGCGAGA 840
 QY 541 CCGCGCGATTTGAGTAGTCAATTCATGCTGTAATACAGCGTGGCGCGCCGACATGAT 600
 Db 841 CCGAGGACATGCTCAATTCCTCTGCTGAGTACAGGCTGCGGCGCGGATTC 900
 QY 601 GAACTGATCGCACCGGAGGATTCGATCGCTGATCCCAAGGAAGATTGTGACT 660
 Db 901 GAACTGATCGCGGAGGATTCACACGCGATTTCCCGCAGGGCGCTTCGTCACC 960
 QY 661 AATGATCCGATGAGCTGTGTGCTGCTGAGCGGCTGCGGATTCGCTACGCTGCGG 720
 Db 961 AACGACTCTCCACCAATGATCGTTGGCTGAAAAACGCGCGCGGATTCGCTACGCGCC 1020
 QY 721 CTGATGCTGATCAACGAGATCAATCTGCGGAGCTGAGATCTCTGCTGCGCGCTTAC 780
 Db 1021 CTGATGCTGATCAACGAGATCAACCGCGGAGATCGATCTGTTCAAAGCTAC 1080
 QY 781 CACTCAGATCCAGCCCGGTTATGCTTATATACCGAAAAAGATGATGCGCTGAGAG 840
 Db 1081 CATTGCTGCGCGCGGATCTACGCGCTTATACCGAGAGCAAAATGCGCTGAGAG 1140
 QY 841 GTACAGTCTGATCAACTCGCTGAGGATTTTGTGAGTCTGCTGCTGCTGCTTAC 900
 Db 1141 GTACAGTCTGATCAATTTTAACTGATTTTCGACGCTGAGCGGCTGCTATCAG 1200
 QY 901 G 901

Db 1201 G 1201

Search completed: December 19, 2004, 01:09:49
Job time : 4562.99 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:46:30 ; Search time 610.159 Seconds
(without alignments)
8001.134 Million cell updates/sec

Title: US-10-759-889-1

Perfect score: 930
Sequence: 1 atgggaacgactaaacgcgat.....ggcgcgggaagagaagtaa 930

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	853.6	91.8	1742	5	AAS89852
2	621.4	66.8	933	11	ABD00728
3	407.6	43.8	1120	5	AAS93346
4	407.6	43.8	1575	5	AAS94105
5	369	39.7	750	5	AAS89851
6	122	13.1	948	11	ABD02039
7	121.8	13.1	1026	11	ABD06582
8	112.6	12.1	927	11	ACH96283
9	112.4	12.1	939	8	ACA24126
10	110.4	11.9	957	11	ABD03348
11	110.4	11.9	1047	11	ABD05418
12	110.4	11.9	1089	11	ABD05289
13	109.6	11.8	945	8	ACA44271
14	108.6	11.7	915	11	ACH96991
15	108	11.6	912	11	ABD02170
16	107.6	11.6	1017	11	ABD13371
17	107.6	11.6	1104	11	ABD13302
18	106.4	11.4	1029	11	ABD15082
19	106.4	11.4	1071	11	ABD14855
20	106.4	11.4	1809	11	ABD15283
21	105.6	11.4	936	8	ACA25126

22	105	11.3	900	3	AAS23355
23	103.6	11.1	909	10	ABZ41110
24	103.6	11.1	912	3	AAS23353
25	103.4	11.1	888	10	ADC19769
26	103.4	11.1	900	3	AAS23354
27	103.4	11.1	1132	10	ADC19786
28	103.4	11.1	47475	3	AAA81465
29	103.4	11.1	110000	3	AAA81489
30	103.4	11.1	349980	3	AAF21612
31	103.2	11.1	933	11	ACH94659
32	102.8	11.1	777	11	ABD06646
33	102.2	11.0	591	10	ADH82126
34	101.4	10.9	915	11	ABD16364
35	101.4	10.9	1011	11	ABD16178
36	100.4	10.8	1062	11	ABD10482
37	100.2	10.8	1008	11	ACH97551
38	100	10.8	100	8	ACD69195
39	100	10.8	100	8	ACD69193
40	100	10.8	100	8	ACD89134
41	99.8	10.7	918	8	ABT14599
42	99.2	10.7	72750	3	AAA81468
43	99.2	10.7	110000	3	AAA81490
44	99.2	10.7	349980	3	AAF21544
45	98.6	10.6	792	11	ABD13408

ALIGNMENTS

RESULT 1

AAS89852

ID AAS89852 standard; cDNA; 1742 BP.

AC AAS89852;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #25656.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008531.

XX

PR 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

PR (HYSE-) HYSEQ INC.

 XX | FA | XX | PI | XX | DR | XX | DR | XX | PT | XX | PS | XX | CC | XX | CC | XX | CC | XX | CC | XX | CC | XX | CC | XX | CC | XX |

Claim 1; SEQ ID NO 25656; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal

Db	184	CGAGGTAATAATCTATTACACGGGCTGCCGTAGGATGCTGTTTGAAGCGCAGGATGTCCAT	243
Qy	241	GAGCAACTGTATGCTTCAATAACACCCCACTCGGAGACCTACGCAATGGCTGTCTTCTCA	300
Db	244	GAGCAATTAACGCAATTAACCAATACCCCACTCGGAGCGCTGCGCATCGGCTGCTCTTCA	303
Qy	301	ACTATGGCACAAATGTTCTCGCCGGGTGACAGCCAAATATGCTGAAGAAATACCCAGGT	360
Db	304	ACTATGGCACAGAATGTTCTGCGCGCATAAACGGCAAAATGCTCAAAGAAATACCCCGGG	363
Qy	361	TTGAGCGTCAATCTGTTTACCCGAATTCACGCCCCGACCTGATTGCGGACGGTCTGGAT	420
Db	364	CTCTC-GTTAACTGTGTACCGGCATACCTGGCCAGACCTGATCGCGACGGCCTTGAC	423
Qy	421	GTGGTANTCGCGTGGCGCGTTCAGAGATCCAGCCTGTTTCCCCCGCTCTGGGCGCG	480
Db	424	GTGGTANTCGCGTGGCGCGCTTCAGAGATCCAGCCTTTTTTCCCGCGCTCTGGGCTCA	483
Qy	481	ATGCCAATGTTGTGCGCCGCGAAAAGCTATCTCACAAATACGGCATACCGGAAAAA	540
Db	484	ATGCCAATGGTATTATGCGCCCTAAAAGCTATCTGGCGCAGGCTGGCAATCCAGAGAA	543
Qy	541	CCGCCGATTTAGTATAGTCAATCATGGCTTGAATACAGCGTGGCGGCCCGCAATGAATTT	600
Db	544	CCGCCGACCTTGCCGGAATGCTGGCTGGAATATAGCGTGGCGCGCGGACATGAGTTT	603
Qy	601	GAACCTGATCGCACCGGAGGATCTCGACTCGCCTGATCCCAAGGAAGATTTGTGACT	660
Db	604	GTCAATATCGCGCGGAGGGATCTCGACCCCGCTGACCCCGCACAGGGGCGCTTCGTGACC	663
Qy	661	AATGATCCGATCAGCTGTGCGCTGGCTGACGGCGGTGCGGGATCGGCTACGTGCGG	720
Db	664	AACGACCTATGACGCTGTGGCTGGCTGACCCGCGGGTGGGTATCGCTATGTACCC	723
Qy	721	CTGATGTGGTGATCAACAGATCAATCGTGGGAGCTGGAGATCTCTGTCTGCGCGTTAC	780
Db	724	TTAAGTGGCGCATCGAAGAGATCAACCGCGTGAACCTTGAGATCTCTGTCTGCCAGTAT	783
Qy	781	CAGTCAGATCCACGCCCGGTTTTATCGCTTATATACCGAAAAAGATGAAGCTGCCGCTGAAG	840
Db	784	CAGTCGATCCGACCCGCTTACGCCCTGTATACGAAAAGGATAACTGCCGCTGAAA	843
Qy	841	GTACAGTCTGTATCAACTCGCTGACGATTATTTTGTGAGTCTGGTAAATGTTTTCAG	900
Db	844	GTGCAGTCTGTATCAACTATCTACGCGAGTATTTTGTGACGCTGGCGAAGATCTATCAG	903
Qy	901	GAGATGCACGGCGCGGA	919
Db	904	GGCATGCACGGCGCGGA	922
RESULT 3			
AAS93346			
ID	AAS93346 standard; cDNA; 1120 BP.		
XX			
AC	AAS93346;		
XX			
DT	13-FEB-2002 (first entry)		
XX			
DE	DNA encoding novel human diagnostic protein #29150.		
XX			
KN	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
XX	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
OS	Homo sapiens.		
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001; 2001WO-US008631.		
XX			
PR	31-MAR-2000; 2000US-00540217.		

PR	23-AUG-2000; 2000US-00649167.
XX	(HYSE-) HYSEQ INC.
PA	Drmanac RT, Liu C, Tang YT;
XX	
PI	WPT; 2001-639362/73.
XX	P-PDSB; ABG29159.
DR	New isolated polynucleotide a
PT	diagnostics, forensics, gene
PT	responsible for genetic disor
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 29150; 1030
XX	
CC	The invention relates to isol
CC	sequences. (I) is useful as h
CC	reaction (PCR) primers, olig
CC	and in recombinant production
CC	in diagnostics as expressed s
CC	genes. (I) is useful in gene
CC	activity of (II) or to treat
CC	useful for generating antibod
CC	polypeptide in tissue, as mol
CC	supplement. (II) and its bind
CC	of sites expressing (II). (I)
CC	involving aberrant protein ex
CC	polypeptide and polynucleotid
CC	diagnostics, forensics, gene
CC	responsible for genetic disor
CC	and to produce other types of
CC	amino acid sequences. AAS6419
CC	coding sequences of the inven
CC	patent did not appear in the
CC	electronic format directly fr
CC	ftp.wipo.int/pub/published_pc
XX	
SQ	Sequence 1120 BP; 271 A; 279
Query Match	
Best Local Similarity 43.8%;	
Matches 416; Conservative	
Qy	481 ATGCCAATGGTGGTGTGGCGC
Db	1 ATGCCAATGGTGGTGTGGCGC
Qy	541 CCGCCCGATTGGAGTAGTCA
Db	61 CCGCCCGATTGGAGTAGTCA
Qy	601 GAATGATCGCACCGGAGGA
Db	121 GAATGATCGCACCGGAGGA
Qy	661 AATGATCCGATCAGCGCTGGT
Db	181 AATGATCCGATCAGCGCTGGT
Qy	721 CTGATGTGGGTGATCAACGM
Db	241 CTGATGTGGGTGATCAACGM
Qy	781 CAGTCAGATCCACGCCCGGT
Db	301 CAGTCAGATCCACGCCCGGT
Qy	841 GTACAGTCGTGATCAACTC
Db	361 GTACAGTCGTGATCAACTC
Qy	901 GAGATGCACG 910

PR	23-AUG-2000; 2000US-00649167.
XX	(HYSE-) HYSEQ INC.
PA	Dzmanac RT, Liu C, Tang YT;
XX	
PI	WPT; 2001-639362/73.
XX	P-FSDB; ABG29159.
DR	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PS	Claim 1; SEQ ID NO 29150; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II), and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 1120 BP; 271 A; 279 C; 323 G; 247 T; 0 U; 0 Other;
	Query Match 43.8%; Score 407.6; DB 5; Length 1120;
	Best Local Similarity 96.7%; Pred. No. 2.7e-115;
	Matches 416; Conservative 0; Mismatches 14; Indels 0; Gaps 0
Qy	481 ATGCCAAATGGTGTGGTCGCGCGAAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
Db	1 ATGCCAATGGTGTGGTCGCGCGAAAAGCTATCTCACAAATACGGCATACCGGAAAAA 60
Qy	541 CCGCGCGATTGTAGTAGTAATCAATGGCTTGAAATACAGCTGCGGCCCGCAGCAATGAATTT 600
Db	61 CCGCGCGATTGTAGTAGTAATCAATGGCTTGAAATACAGCTGCGGCCCGCAGCAATGAATTT 120
Qy	601 GAACATGATCGCACCGGAAGGATCTGCACTCGGCTGATCCACAAGAGAGATTTGTGACT 660
Db	121 GAACATGATCGCACCGGAAGGATCTGCACTCGGCTGATCCACAAGAGAGATTTGTGACT 180
Qy	661 AATGATCCGATGACGCTGGTGGCTGCTGACCGCGGTCGCCGGATCGCCTACGTCGCG 720
Db	181 AATGATCCGATGACGCTGGTGGCTGCTGACCGCGGTCGCCGGATCGCCTACGTCGCG 240
Qy	721 CTGATGTCGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCGCGGTAC 780
Db	241 CTGATGTCGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCGCGGTAC 300
Qy	781 CAGTCAGATCCACGCCCGTTTTATGCGTTATATACGGAAAAAGATAAGCTGCGCGTGAAG 840
Db	301 CAGTCAGATCCACGCCCGTTTTATGCGTTATATACGGAAAAAGATAAGCTGCGCGTGAAG 360
Qy	841 GTACAGGTCGTGATCAACTCGCTGACGGATATTTTGTGAGTTCGGTAAATGTTTCAG 900
Db	361 GTACAGGTCGTGATCAACTCGCTGACGGATATTTTGTGAGTTCGGGCGGAGGACCG 420
Qy	901 GAGATGCACG 910

```

Db      421 TTCGTGCACG 430
|||||
RESULT 4
AAS94105
ID AAS94105 standard; cDNA; 1575 BP.
XX
AC AAS94105;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #29909.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG29918.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 29909; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1575 BP; 402 A; 376 C; 459 G; 338 T; 0 U; 0 Other;

Query Match 43.8%; Score 407.6; DB 5; Length 1575;
Best Local Similarity 96.7%; Ident. No. 3.2e-115;
Matches 416; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 481 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAATACGGCATACCGGAAAAA 540
DB 1 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAATACGGCATACCGGAAAAA 60

```

CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 750 BP; 163 A; 197 C; 204 G; 186 T; 0 U; 0 Other;

Query Match 39.7%; Score 369; DB 5; Length 750;

Best Local Similarity 82.7%; Pred. No. 2e-103;

Matches 472; Conservative 0; Mismatches 15; Indels 84; Gaps 1;

```
QY 340 ATGCTGAAGATACCCAGGTTGAGCGTCAATCTGGTTACCGGAATCCAGCCCGAC 399
DB |||
DB 1 ATGCTGAAGATACCCAGGTTGAGCGTCAATCTGGTTACCGGAATCCAGCCCGAC 60
QY 400 CTGATGCCGACGCTGATGATGTTGATCCGCGCGTTCAGGATTCAGGCTG 459
DB |||
DB 61 CTGATGCCGACGCTGATGATGTTGATCCGCGCGTTCAGGATTCAGGCTG 120
QY 460 TTTTCCCGCGTCTGGCGCGATGCCAATGTTGGTGTGCGCGCGAAAGCTATCTACA 519
DB |||
DB 121 TTTTCCCGCGTCTGGCGCGATGCCAATGTTGGT----- 155
QY 520 CAATACGGATACCGGAAACCCCGGATTTGAGTAGTCATTCGTTGATACAGC 579
DB |||
DB 156 -----C 156
QY 580 GTGCGGCCGACAAATGAATTTGAATCTGATCGACCGGAAGGATCTCGACTCGCTGATC 639
DB |||
DB 157 GTGCGGCCGACAAATGAATTTGATCTGATCGACCGGAAGGATCTCGACTCGCTGATC 216
QY 640 CCACAGGAAGATTGTGACTAATGATCGATGACGCTGGTGGCTGCTGACGCGGGT 699
DB |||
DB 217 CCACAGGAAGATTGTGACTAATGATCGATGACGCTGGTGGCTGCTGACGCGGGT 276
QY 700 GCGGGATCGCTACGTGCGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTG 759
DB |||
DB 277 GCGGGATCGCTACGTGCGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTG 336
QY 760 GAGATCCTGCTGCGCGTTACAGTCAGATCCACCGCCGGTTTATGCGTTATATACCGAA 819
DB |||
DB 337 GAGATCCTGCTGCGCGTTACAGTCAGATCCACCGCCGGTTTATGCGTTATATACCGAA 396
QY 820 AAGATACGCTCGCTGAGGTACAGTCTGATCACTGCTGACCGATATTTTGT 879
DB |||
DB 397 AAGATACGCTCGCTGAGGTACAGTCTGATCACTGCTGACCGATATTTTGT 456
QY 880 GAGTGGTAAATTTTTCAGGAGATGCAACG 910
DB |||
DB 457 GAGTGGGCGGAGGACCGTTCGTGACG 487
```

RESULT 6

ABD02039

ID ABD02039 standard; DNA; 948 BP.

AC ABD02039;

XX

XX

DT 29-JUL-2004 (first entry)

XX

DE Pseudomonas aeruginosa polynucleotide #643.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
FN US551795-B1.
XX
PD 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
FI
XX WPI; 2003-615309/58.
XX
XX P-PSDB; ABO68468.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 643; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 948 BP; 156 A; 314 C; 330 G; 148 T; 0 U; 0 Other;

Query Match

Best Local Similarity 13.1%; Score 122; DB 11; Length 948;

Matches 417; Conservative 0; Mismatches 470; Indels 3; Gaps 1;

```
QY 1 ATGGAACGACTAAACGCGATGTCGGTGTTCGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
DB |||
DB 40 ATGGAATTCGCTGAAGGCGATGCGGATCTTCGCCACTGTGTGGACAAGGCTCGATGGCA 99
QY 61 GCGCGCCCGCAGACAGCTACAGATGAGGGTTTCGTCATCGTCAGACGGTATCAAACTG 120
DB |||
DB 100 GCGCGCCCGCAGACAGCTGCGGATGACCCCTTCGGCAGTTCAGCCAGAGATCCGCAAGCTG 159
QY 121 GAAGATGAGTTGAGGTAAAGCTGTAAACCGTAGCACACGCGAGCATTTGGCTTCGCCGAA 180
DB |||
DB 160 GAGACCGCGCCCGCAGGTCACCTTGTGTCATCGCACACCGCCGCGGTGACCTGACCGAG 219
QY 181 GCGGTAGAAATTTACTACAGGGCTGCGGTTCGATGCTTCATGAAGTCAGAGATGTTTAT 240
DB |||
DB 220 GCGCGCGAGGCGTTCTATCGCAGTTGCGGCGCAGATCTGGCGATCGCCGAGGAGCGGAG 279
QY 241 GAGCACTGTATGCTTCAATTAACACCCCATCGGACGCTACGCAATTTGCTGCTTTCTTCA 300
DB |||
DB 280 CGCGCGCTTGGCGAATGGCGCGATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 339
QY 301 ACTATGSCACAAAATGTTCTCGCCGGGTGACACCCAAAATGCTGAAAGAAATACCCAGGT 360
DB |||
```

Db 340 GGGTTCTCGGCACGCTGATCACCCAGCGCTGAAACCGTTGCTGAGAACACCCGCGAG 399
QY 361 TTGAGCGTCAATCTCGTTACCGGAATTCAGCCCCCGGACCTGATTTGCCGAGCGTCTGGAT 420
Db 400 TTGCGCGCTGACGCTGTTCTTCAGGACGAGCGATCGATCTGGTGGCCGAGCGATCGAC 459
QY 421 GTGGTATCCGGTGGCGGCTTGCAGGATTCACGCTGTTTCCCGCGCTCTGGCGCG 480
Db 460 CTGGCGATACGGTGGCGAATCTCCCGGACTCCAGCTGGTGGCGGCGACCTCGGCGAT 519
QY 481 ATGCCAATGGTGTGGCGCGGAAAGCTATCTCACAAATAGCGATACCGGAAAAA 540
Db 520 TGGAGCAGCGTCTCTGCGCGCGCGGCGCTATTTGGCGCAACGCGCGCCCATCAACCGT 579
QY 541 CCGCGGATTTGAGTAGTCAATTCATGGCTT---GAATACGCGTGGCGCGCGAGATGAA 597
Db 580 CCGGAGCATTTACCGAGGTGAGTGGATTTCCCTGAAACCTCGAACCACTCAACAC 639
QY 598 TTGAACTGATCGCACCGGAAGGATCTCGACTCGCCTGATCCCAAGGAAGATTGTG 657
Db 640 CTGACGCTCAGCGGCGCGGCGGGAAGTCTGCAAACTGGCGCTGGAGCCACGGGTGCC 699
QY 658 ACTAATGATCCGATGACGCTGTGTGGCTGAGCGCGGGTGGCGGATCGCTTACGTG 717
Db 700 GCAACGGGATCTCGCGGTGCGTCACTTCACCCCTCGACGACTCGCGGTGTCTTACCG 759
QY 718 CGCTGATGTGGTGTATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGTGCGCGGT 777
Db 760 CGCTGCCGAGGTGCGGCGCGCTGACGAGCGGCGCTGCAACAGCTATGCGGAG 819
QY 778 TACAGTCAATCCAGCGCGGTTTATGGTTATATACGAAAGATAGTGGCGGTG 837
Db 820 TGGAGGATTCGCGGGTCTGCGCATCTATGCGGTGACGCCACGCGGCGGCGCGCC 879
QY 838 AAGGTACAGTCTGATCACTACCTCGCTGACGGATTATTTGTTGAGGTGCG 887
Db 880 AAGGTGAAGTCCCATGAGGCGCTGCGGAGCGTTGCGGACGAGCGG 929

RESULT 7

ID ABD06582 standard; DNA; 1026 BP.
XX AC ABD06582;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polynucleotide #5186.
XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX KW antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX WPI; 2003-615309/58..
XX DR P-PSDB; AB073011.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 5186; 455pp; English.
XX PS The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX SQ

Sequence 1026 BP; 152 A; 384 C; 339 G; 151 T; 0 U; 0 Other;

Query Match 13.1%; Score 121.8; DB 11; Length 1026;
Best Local Similarity 47.1%; Pred. No. 8.9e-27;
Matches 409; Conservative 0; Mismatches 457; Indels 3; Gaps 1;

QY 1 ATCGAAGCACTAAACGCAATGCGGTGTTTGGCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 106 ATGGACGCGCTGACCGCCACTCGCGTCTTCGTGCAAGTCAATGACAGCGCGACGACG 165
QY 61 GCCCGCGCAGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGAGCGTATCAAACTG 120
Db 166 GCCCGCGCAGCGCTTGGACATGTCGGGGGATGTTCTCGGCTACCTCCCGGATG 225
QY 121 GAAGATGAGTTGAGGTAAGCTTTAAACCGTAGCACAGCAGCATTTGGCTGACCGAA 180
Db 226 GAGGATGGGTGGCGCAGCGCTGTGTCACCGCAGCACCGCGCTGAGCTGACCGAC 285
QY 181 GCCGCTAGAAATTTACAGGCTCGCTGATGCTTCAATGAGTGCAGATGTTCTAT 240
Db 286 GCCCGCGCAGCGTGTGCGCAGTCTCGCGAGATGCTCGCGTGGCGCGACCGCTGACG 345
QY 241 GAGCAACTGTATGCTTCAATTAACACCCCATCGGACCGCTACGATTTGGCTGTTTCA 300
Db 346 GCCACGGCGCAGACCGCGCGCAGCGCGCGGCGACTCTGGCATCACGACGACGCTG 405
QY 301 ACTATGCAAAATGTTTTCGCGGCTGACAGCCAAATGCTGAAAGATACCGAGT 360
Db 406 TCCTTGGCGCAGCGCTGGGTGACCCGTGCGCGCGCGCTTCTGAGCGCTACCGCGGC 465
QY 361 TTGAGCGTCAATCTGTACCGGAATTTCCAGCCCGCGACTGATTTCCGACGCTCTGGAT 420
Db 466 ACCGCCATCGACCTGAGGTCAACACCGAGTGGTCAACTGTTGAGGAGCGCATCGAC 525
QY 421 GTGGTATCCGCTCGCGCGCTTGCAGGATTCAGGCTGTTTTCGCGCTCTGGCGCGG 480
Db 526 CTAGCCCTGGCATCGCAACACGCTCGACCCCAACCTGATCGCGCGCGCTCTCGCGAA 585
QY 481 ATGCCAATGGTGTGGCGCGGAAAGCTATCTCACAAATACCGCATACCGGAAAAA 540
Db 586 TGGCGCTCGGTGATCTGCGCGCGCGCGGACTACCTGCGCGCGCACCGCGCGCGCG 645
QY 541 CCGCGCGGATTTGAGTAGTCAATTCATGGCTTGAATACAGGCTGGCGCGCGCAATGTT 600
Db 646 CCGGAGGACCTGGCGTGCACACCTGCTGACCTACTCTTCTGCGCGCGCTCTC 702
QY 601 GAAGTATCGACCGGAAGGATCTGACTCGCTGATCCACAAAGAGATTTGTGACT 660
Db 703 TGGCAGTTTCAGCGCGACGCGCGGACCGCGCTGCGGTGGGTGGCGCTCTCAGCGCC 762
QY 661 AATGATCCGATGACGCTGGTGGCTGCGTGAACCGCGCGGTGCGGGATCGCTACGTCGCG 720
Db 763 AAGCATCCACCGTGTGCTGGAGCGCGCGCGCGCGGTGCGGGATCAGCAGCAGCGCG 822

QY 721 CTGATGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTGTCGGCGTTAC 780
DB 823 CTGTACTCGGACACCCCGCTGATCCGTAGCGGACGCTGCTGGCGCTGCTGCGAGTGG 882
QY 781 CAGTCAGATCCAGCCCGGTTTATGCGTTATATACGAAAGATAAGCTGCGCTGAAG 840
DB 883 TCGCCGAGGTGCTCGGATCCACGGGTGACGCTCGCGCGGAGATGCGCGCGG 942
QY 841 GTACAGGTGCTGATCAACTGCTGACGGA 869
DB 943 CTGCGGCGCTGCTGACCTCTCTGTCGA 971

RESULT 8

ACH96283
ID ACH96283 standard; DNA; 927 BP.

AC ACH96283;

DT 29-JUL-2004 (first entry)

DE Klebsiella pneumoniae polynucleotide seqid 2078.

XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

OS Klebsiella pneumoniae.

XX US6610836-B1.

PN 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX P-PSDB; ABO62732.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 2078; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

XX SQ Sequence 927 BP; 150 A; 278 C; 335 G; 164 T; 0 U; 0 Other;

XX Query Match 12.1%; Score 112.6; DB 11; Length 927;

XX Best Local Similarity 45.7%; Pred No. 56-24; Indels 0; Gaps 0;

XX Matches 394; Conservative 0; Mismatches 469;

QY 1 ATGGACGACTAAACGATGCGGTGTTGCAAGTAGTGAATTTGGCTTTTACC 60

DB 16 ATGGACAAATTACGAGGATGAGACGTTATTGCGGTGGTGGAGCTTTTACC 75

QY 61 GCGCGCGGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAACTG 120

DB 76 GGGCGGCGTCCGCGCTGCGGCTGTCGGGGTATGTTGGGAATATATCGCCAGCTG 135

QY 121 GAAGATGATTGACGTAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTGACCGAA 180

DB 136 GAAAGCCAGCTGGCGAGCGCGTCTCTGGAGCGTAACACCCGACGCCAGAGCTCACCGAC 195
QY 181 GCCGTAAGATTTACTACAGGCGTGGCGTCTGATGCTTCATGAAGTGCAGAGTTCAT 240
DB 196 GCCGGCGAGTCTATTTTACGAAGCAAGCGGTATGAGCAGGTCTCCATGCCGAA 255
QY 241 GAGCAACTGATGCTTCAATAACACCCCATCGGACGCTAGCAGTATGGCTGTTTCA 300
DB 256 AGCGCGTTCGAGCGGCTGCGGCTGGCGCGCGGACCCCTGGGGTGAGCGCTCCACC 315
QY 301 ACTATGACACAAATGTTCTGCGCGGCTGACAGCCCAAAATGCTGAAGAATACCCAGGT 360
DB 316 TCGTTCCGGCCAGCGTATCGCCCGCTACGCCCAACTTCTTCAGAGGCTGGCGTGGC 375
QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCGCCGACCTGATTCGCGACGCTCTGGAT 420
DB 376 GTCCGGGTGGAACCTGGATCTGACTAACCGGATGTCGATCTGGTGGATGAGGGGTTCGAT 435
QY 421 GTGGTATCCGCTCGCGCGGTTCAGGATTCAGCCTGTTTCCCGCCGCTCTGGGCGCG 480
DB 436 CTGGCGATCCGATATCGCGGAGATCCATCAGAGGATCTGGTGGCTCGCTACCTGGCGCG 495
QY 481 ATGCCAATGGTGGTGTGCGCGCGGAAAGCTATCTACAAATACGGCATACCGGAAAAA 540
DB 496 TATCGGATGATGATCTGCGCGCGCGCGCTTATCTGGCGCTTACGGGACCCCGGACG 555
QY 541 CCGCCGATTTGAGTACTCATTTATGCTTTGAATACAGCTGCGGCCCGCAATGAATTT 600
DB 556 CCAGAGGATCTCGCGGATCATCTCTGCTCTGCCATACGCTTGGACCGCGCAACGAG 615
QY 601 GNACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTGTGACT 660
DB 616 TGGCGGCTCGCGGCTGAGGAGAGTGGTGGTGAACGCGATGCGGTTTTGCGGTGC 675
QY 661 AATGATCCGATGACGCTGGTGGCTGGCTGACGCGGGTCCCGGATCGGCTACGTGCGG 720
DB 676 AACGATGGCTACGCGCTGCGCGAGCGGCGATCGCGGGCGCGGCTGCTGATGACGCGG 735
QY 721 CTGATGTTGGTATCAGCAGATCAATCGTGGGAGCTGAGATCTGCTGCCGCTTAC 780
DB 736 GAGTACTGCTGGCGGATGCGTGGCGAGCGGAGCTGGTTCCGGTGTGGAGCGGTGG 795
QY 781 CAGTCAGATCCAGCCCGGTTTATGGGTTATATACGAAAAAGATAAGCTGCCGCTGAAG 840
DB 796 ACTCCGCGAGCGCGGCTGTCATCTGCTTGGCGCAGGATCGTGGGCCCTTCCGGAAG 855
QY 841 GTACAGTCTGATCAACTCGCT 863
DB 856 CTGACGAGTTTATCGCCACCT 878

RESULT 9

ACA24126
ID ACA24126 standard; DNA; 939 BP.

XX ACA24126;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #5783.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

XX Borrelia cepacia.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US0009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlssen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU20256.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 11996; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 939 BP; 145 A; 319 C; 327 G; 148 T; 0 U; 0 Other;
 SQ

Query Match
 Best Local Similarity = 47.8%; Pred. No. 7e-24;
 Matches 394; Conservative 0; Mismatches 421; Indels 9; Gaps 2;

QY 1 ATGGAACGACTAAACGCGATCGGTTGTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
 Db |||||
 QY 1 ATGGAATCGTTACAAACATGCGGTTATTCGTCGCGTGTGCGACGCGGCGAGCTTTACC 60
 Db |||||
 QY 61 GCGCGCGCCAGACGAGTACAGATGACGCTTTCGTCATCAGTACAGCGGTATCAAAATG 120
 Db |||||
 QY 61 GCGCGCGCCAGACGAGTACAGATGACGCTTTCGTCATCAGTACAGCGGTATCAAAATG 120
 Db |||||
 QY 121 GAAGTGTAGTTCAGTAAAGTGTAAACCTAGCAGACGCGATTCGCTGACCGAA 180
 Db |||||
 QY 121 GAGGCCACCTGCGGACCGGCTTCCTGAAACCGACGACGCGCGGATCCGCTGACCGAG 180
 Db |||||
 QY 181 GCGCGGTAGAAATTTACTACGAGGCTCGCGTCTGATGCTTTCATGAAGTGAGGATTTTAT 240
 Db |||||
 QY 181 GCGCGCGAGGCTTATCTGACGCGCTCGCAACAGATCCTCGCTTACGTCGACCGCGGAA 240
 Db |||||

QY 241 GAGCAACTGTATGCTTCAATTAACACCCCATCGGACGCTACGATTTGGCTTCTTCA 300
 Db |||||
 QY 241 GCGGAGCGGGGACGCGACGCGCGCGCTCGCGCAAGCTGAAGTTCATTGCTTCAAG 300
 Db |||||
 QY 301 ACTATGGCACAATAATGTTCTCGCCGGGCTGACAGCCAAATCTGAAAGAAATACCCAGT 360
 Db |||||
 QY 301 AGCTCGGCGACACATATCTGTCGCGCCATCGCGGCTACCGGAGGCTATCCGAC 360
 Db |||||
 QY 361 TTGAGCGTCAATCTGCTTACCGAAATTCAGCCCCGACCTGATTCGCCAGCGTGTGAT 420
 Db |||||
 QY 361 GTGACGTCGAGCTGACGCTCGCGCAGCGGATGCGCGACCTGCTCGACGAGGCTACGAC 420
 Db |||||
 QY 421 GTGCTGATCGGCTCG ---GCGGTTGAGGATTCAGGCTGTTTCCCGCGCTCTGGGC 477
 Db |||||
 QY 421 GTGCGATCGTCTGCGCGCGACCTGCGCGGATTCGGGCTCGTGTGAGCGGCTCGGC 480
 Db |||||
 QY 478 GCGATGCCAATGTGTGTCGCGCGGAAAGCTATCTCACAAATACGCGATACCGAA 537
 Db |||||
 QY 481 GAGAGCTACAGGCTGCTGCGCATCGCGGCTACGTCGAGTCGCGACGCGCTGCGCAG 540
 Db |||||
 QY 538 AAACCCCGGATTTGAGTAGTCAATTCATGCTTGAATACAGCTGCGG-----CCGAC 591
 Db |||||
 QY 541 CGCGCGCGGATCTCCACAGACAGCTGTCCTCGGGATGTCGCGCGCGGCTTTCATTC 600
 Db |||||
 QY 592 AATGAATTTGAATGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGAAAGA 651
 Db |||||
 QY 601 GAGCAATGGCGCTGCGCGGCGCGAACGCGGAGGTCGTCGCGATCACGCGCGCGCG 660
 Db |||||
 QY 652 TTTGTGACTAATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
 Db |||||
 QY 661 TTTGCGGTGAACGTCGCGGAGGCGCTCGCGGTGCGCGGAGGAGTGGCGCTGCGC 720
 Db |||||
 QY 712 TACGTCGCGCTGATGCGGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTG 771
 Db |||||
 QY 721 GGGCTGCGGCTCTATTTCGCGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Db |||||
 QY 772 CCGCGTTTACGATCAGATCCAGCCCGGTTTATGCGTTATATAC 815
 Db |||||
 QY 781 CCCGAGTACCGTCCGACGATGATGAACATCTACGCGCTGTATCC 824
 Db |||||

RESULT 10
 ABD05348
 ID ABD05348 standard; DNA; 957 BP.
 AC ABD05348;
 XX
 XX 29-JUL-2004 (first entry)
 DT
 XX
 XX Pseudomonas aeruginosa polynucleotide #3952.
 DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.
 XX
 XX Pseudomonas aeruginosa.
 OS
 XX
 XX US551795-B1.
 FN
 XX
 XX 22-APR-2003.
 PD
 XX
 XX 18-FEB-1999; 99US-00252991.
 PF
 XX
 XX 18-FEB-1998; 98US-0074788P.
 PR
 XX 27-JUL-1998; 98US-0094190P.
 PR
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX
 XX Rubenfield MJ, Nollig J, Deloughery C, Bush D;
 PI
 XX WPI; 2003-615309/58.
 DR
 XX P-PSDB; ABO1777.
 DR
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT

PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 3952; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biotech technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX Sequence 957 BP; 145 A; 341 C; 327 G; 144 T; 0 U; 0 Other;

Query Match 11.9%; Score 110.4; DB 11; Length 957;

Best Local Similarity 50.0%; Pred. No. 2.9e-23;

Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAAACGACTAAACGCTATGCTGGTGTGGCAAGTAGTGAATTTGGCTCTTTTACC 60
DB 34 ATGGACCGTTTCATCGCATGAGGCGCTTTGCCGCTGGTGGATACCGGACGCTTACC 93
QY 61 GCGCGCGCAGACAGTACAGTACAGCTTGTTCATCAGTACAGCTATCAAACTG 120
DB 94 AAGCGCGCGAGACGCTGCATATGAGCCGACGCGTCCACGCTGTCGAGAACTG 153
QY 121 GAAGATGAGTTGCGAGTAAAGCTGTAAACCGTAGCACACGAGCATTTGGCTGACCGAA 180
DB 154 GAAGCGCGCTGGGTAGCGCTGCTCAACCGCACACCGGCAAGCTCACTTACCGCC 213
QY 181 GCGGTAGAAATTTACTACAGGCTGCGCTGATGCTTTCATGAAGTGCAGGATTTTCA 240
DB 214 GACGCGCGCGCTACTACAGCGGGTGTGACGCTGCTGGCGAACTGGACGATCCGAG 273
QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGACGCTACGATGGCTGTTCTTCA 300
DB 274 ACCAGCTTTCAGCGCCCGCTTGGCGCCCGCGGCTCTCCGATGACGTACCGAGC 333
QY 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGT 360
DB 334 CCGCTGGCGGGTGTGCTGATCCGCGGTTGCGCGAGTTCATGCGGCTATCCGAG 393
QY 361 TTGAGCGTCAATCTGTACCGAATTCAGCGCCCGACCTGATTCGCGAGGCTTGGAT 420
DB 394 ATCCAGATGACCTGGCGGTGAGCGATCGGCTCGTGGCTGATCGAGAACTCGAC 453
QY 421 GTGGTATCGCGGTGCGCGGTGACAGGATTCAGGCTGTTTCCGCGCTCTGGCGCG 480
DB 454 TCGTATCGCGCGCGGAGTGTGCGGACAGTGTGATAGCGCGGGTTCGCGAC 513
QY 481 ATGCCAATGGTGTGCGCGGAAAGCTATCTACACAATAAGGATACCGGAAATA 540
DB 514 CTGCACTGGGGTCTACCGCGCACCGGCTACTTGCAGCGCGCGGGTTCGCGAGCCAC 573
QY 541 CCGCGCGATTG 552
DB 574 CCGCGGAACTG 585

RESULT 11

ABD05418

ID ABD05418 standard; DNA; 1047 BP.

XX

AC

XX ABD05418;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polynucleotide #4022.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.

OS Pseudomonas aeruginosa.

PN US5551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

XX P-PSDB; A5071847.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 4022; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biotech technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX Sequence 1047 BP; 167 A; 361 C; 355 G; 164 T; 0 U; 0 Other;

Query Match 11.9%; Score 110.4; DB 11; Length 1047;

Best Local Similarity 50.0%; Pred. No. 3.1e-23;

Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCTATGCTGGTGTGGCAAGTAGTGAATTTGGCTCTTTTACC 60

DB 44 ATGGAACGCTTTCATCGCATGAGGCGCTTTGCCGCTGGTGGATACCGGACGCTTACC 103

QY 61 GCGCGCGCAGACAGCTACAGATGAGGCTTTCGTTCATCATCATGAGCGGTATCAAACTG 120

DB 104 AAGCGCGCGAGACGCTGTCATGAGCGCGCACCGGCTACCCAGTGTGGTCCAGCACTG 163

QY 121 GAAGATGAGTTGAGGTAAGCTGTAAACCGTAGCACAGGACGATTCGCTGACCGAA 180

DB 164 GAAGCGCGCTGCGGGTACCGCTTGTCAACCGCACCGGCAAGCTCAACCTGACCGCC 223

QY 181 GCGGTAGAAATTTACTACAGGCGTCCGCTCGTATGCTTCATGAAGTGCAGGATGTTTCA 240

DB 224 GACGCGCGCGCTTACTACGAGCGGCTGTACGCTCTGCTGGCGAAGTGGACGATGCCGAG 283

QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGAGCGCTACGCTATGCTGTTCTTCA 300

Db 284 ACCAGGCTCTCCAGCGCCGCTTGGCGCCGCGGCGTCTCCGATCGACGTACCGAGC 343
Qy 301 ACTATGGCACAATATGTTCTCGCGGGCTGACAGCCAAATCTGAAAGAAATACCCAGT 360
Db 344 CCCTGGCGGGTGTCTGATCCCGGGTTCGCGAGTTCCATGCGGCTATCCGGAG 403
Qy 361 TTGAGCGTCAATCTGTTTACCGGAATCCAGCCCGCGACCTGATCCGACCGTCTGGAT 420
Db 404 ATCCAGATCGACCTGGCGGTGAGCGATCGCTCGTACCTGATCGACGAGAACTCGAC 463
Qy 421 GTGGTATCCGCTCGCGGGTTCAGGATCCAGCCTGTTTTCGCGCGTCTGGCGCG 480
Db 464 TGGGTATCCGCGCGCGGAACTGCGCGACCTGCTGATAGCGCGGGTCTGGCGAC 523
Qy 481 ATGCCAATGTTGTGCGCGCGGAAAGCTATCTCACAAATAGCGCATACCGGAAAAA 540
Db 524 CTGCAACTGGGGTCTACGCGGCAACCGGCTACTTGCAGCGCGCGGCTGCCAGCCAC 583
Qy 541 CCGCGCGATTG 552
Db 584 CCGCGCGAACTG 595

RESULT 12

ABD05289/c

ID ABD05289 standard; DNA; 1089 BP.

XX AC ABD05289;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polynucleotide #3893.

XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;

XX KW antibacterial.

XX OS Pseudomonas aeruginosa.

XX PN US651795-B1.

XX XX 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX XX (GENO-) GENOME THERAPEUTICS CORP.

XX XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX DR P-PSDB; ABO71718.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 3893; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABD01397-

CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 1089 BP; 168 A; 372 C; 372 G; 177 T; 0 U; 0 Other;

Query Match

Best Local Similarity 11.9%; Score 110.4; DB 11; Length 1089;

Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

Qy 1 ATGGAACGACTAAACGCAATGCGGTGTTTCCAAAGTAGTTGAAATTTGGCTCTTTTACC 60

Db 981 ATGACCGTTTCGATGCGATGAGGCTTTGCGGTGTTGCGGTGATACCGGACGCTTACC 922

Qy 61 GCGCCGCCAGACAGCTACAGATGAGGTTTCTCATCATGTCAGAGGTTATCAAACTG 120

Db 921 AAGCGCGGAGAGCGTGCACATGAGCGCGCACAGGTTCAACCCAGTTGGTCCAGCAACTG 862

Qy 121 GAGATGAGTTGAGGTAAAGCTGTTAAACGTTAGCACACGAGCATTTGCCCTGACCGAA 180

Db 861 GAGCGCGGCTGCGGTTACGCTTCTCAACCGCACACGCGCAAGCTCAACCTGACCGCC 802

Qy 181 GCCGTAGAATTTACTACCGGGCTGCCGTCTGATGCTTCAAGTGCAGGATGTTTCA 240

Db 801 GACGGCGCGCTACTACGAGCGGTTGTTACGTTCTGTCGCGGAACTGGACGATGCCGAG 742

Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTAGCATTTGGCTGTTCTTCA 300

Db 741 ACCAGCCTCTCCAGCGCGCCCTGCGGCGCCCGCGGGCTCTCCGATCGACGTACCGAG 682

Qy 301 ACTATGACACAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360

Db 681 CCGCTGCGCGGCTGCTGCTGATCCCGGCTTCCCGAGTTCCATGCGCGCTATCCGAG 622

Qy 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCGACCTGATTCGCGACGCTCTGGAT 420

Db 621 ATCCAGATCGACCTGGGCGTGGCGATCGGCTCGTCCGCTGATCGACGAGAACGTCGAC 562

Qy 421 GTGGTATCGCGTCCGCGGCTTGCAGGATTCAGGCTTTTCCGCGCGCTCTGGGCGCG 480

Db 561 TGGCTATCCGCGCGCGAACTGGCGCACGATCGCTGATAGCGCGGGTCTGGCGAC 502

Qy 481 ATGCCAATGTTGTTGCGCGCGGAAAGTATCTTACACAAATACGACATACCGGAAAAA 540

Db 501 CTGCAACTGGGGGTCTACGCGGCAACCGGCGCTACTTGCAGCGCGCGCGCTGCCAGCCAC 442

Qy 541 CCGCGCGATTG 552

Db 441 CCGCGCGAACTG 430

RESULT 13

ACA44271

ID ACA44271 standard; DNA; 945 BP.

XX AC ACA44271;

XX DT 19-JUN-2003 (first entry)

XX DE Prokaryotic essential gene #25928.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX OS Pseudomonas putida.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX XX

PR 21-MAR-2001; 2001US-00815242.
 DR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JB, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR P-PSDB; ABU40401.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 32141; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 5213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing a polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 945 BP; 188 A; 305 C; 277 G; 175 T; 0 U; 0 Other;
 Query Match 11.8%; Score 109.6; DB 8; Length 945;
 Best Local Similarity 50.7%; Pred. No. 5.2e-23;
 Matches 290; Conservative 0; Mismatches 279; Indels 3; Gaps 1;
 QY 1 ATGGACAGCTAAACGCGATGTCGGTGTGGCAAGTAGTTGAATTTGGCTCTTTACC 60
 DB 1 ATGGACACCTGCAAAACATGCTGCTTTTCAGTGGTAGCCCACTGGCAGCTTCACT 60
 QY 61 GCGCGCCGACAGCTACAGTACAGAGCGTTGTCATCAGTCACAGCGTATCAAACTG 120
 DB 61 GCGCGTGGCGGCAACTGATGATGACGACACCGGACGTCGCGGCGGTCTTCAACCTG 120
 QY 121 GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACAGCAGGATTCGCTGACCGAA 180
 DB 121 GAAGCCCATCTGCAACACAGGCTGCTCAACCGTACACCGCGGATTCGCTGACCGAA 180
 QY 181 GCGGTAGAAATTACTACAGGCGTCCGTCGTATGCTTCAATGAAGTCAGGATTTTCAT 240
 DB 181 GCGGCAAGCGTTACTGATGCGTTTGGCAACAGATTTTACCTAGTCGAAAGCCGAG 240

241 GAGCACTGTATGCTTCAATAACACCCCATCGGAGCGTACGCAATGGCTGTCTTCA 300
 241 GCGAAGCCAGCAGCAGCCCATGCGCGCCGCGCAGTTGAAGTGCATTGATGACT 300
 301 ACTATGGCACAAATGTTCTCGCGGGCTGACAGCCCAAAATGCTGAAGATATCCAGGT 360
 301 GGGTGGCCAGCACTTCTGTTGATGCCATCCCGCGCTACCGCGAAACGACCCCGAC 360
 361 TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCGCCGCTGATTGCGAGCGTCTGGAT 420
 361 GTGACCTTCGATTGACCATGSCCAACGCGTGCCTGCTCTCGAGGAGGTTATGAC 420
 421 GTGTGTATC---CGGTGCGCGCGTTGAGAGATTCAGCCCTGTTTCCGCGCTCTGGG 477
 421 GTCTCCATCGTGTGGCCACCGAATCTCCGAGCTCGGGTTCGTTTCGAGCGCTGGGC 480
 478 GCGATGCCAATGTTGGTGGCGCGCAAAAGCTATCTCACACAATACGGCATACCGGAA 537
 481 ATCACTACAGCATTTGTTGCTGCTCGCCGCTACATCGCCAGCATGGGTGGCGAC 540
 538 AAACCCGCGGATTTGAGTAGTCAATTCATGGCT 569
 541 AAGCTGCGGACCTGCTCAAGCAGCGCTGCT 572

RESULT 14
 ACH96991
 ID ACH96991 standard; DNA; 915 BP.
 AC ACH96991;
 XX 29-JUL-2004 (first entry)
 DE Klebsiella pneumoniae polynucleotide seqid 2786.
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
 OS Klebsiella pneumoniae.
 PN US6610836-B1.
 XX 26-AUG-2003.
 PD 27-JAN-2000; 2000US-00489039.
 PF 29-JAN-1999; 99US-0117747P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 PI Breton GL, Osborne M;
 XX
 DR WPI: 2003-895346/82.
 DR P-PSDB; ABO63440.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 2786; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
 CC pneumoniae polypeptide of the invention
 XX
 SQ Sequence 915 BP; 149 A; 276 C; 297 G; 193 T; 0 U; 0 Other;
 Query Match 11.7%; Score 108.6; DB 11; Length 915;
 Best Local Similarity 46.9%; Pred. No. 1e-22;

Matches 412; Conservative 0; Mismatches 459; Indels 8; Gaps 2;

QY 10 CTAACAGCATGTCGGTGTTCGCAAGTAGTGAATTTGGCTCTTTACCGCCGCCGC 69

Db 37 CTTACGGGTGGCCATTTTATCGCGGTGGTAGATGCGCGAGTTTACCTCGCCGCA 96

QY 70 AGACAGTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAACTGGAAGATGAG 129

Db 97 CGCGCGCTGGGAGACACAGGGGTGGTTCAGCTTAACGTCAGCGAGCTGGAAATGAG 156

QY 130 TTCAGGTAAAGCTGTAAACCGTACACAGCAATTTGGCTGACCGAAGCCGGTGA 189

Db 157 CTGGGGTACGCTGCTGCTGCTTCGACCGCGCGTTTCGGCTCACTGACCGCGGGTG 216

QY 190 ATTACTACAGCGCTGCGCTGTATGCTTTCATGAAGTGCAGATGTTTCATGACAACTG 249

Db 217 CTGTTTATCAGCGCGGGTGGCTCTTGAACGCGCGGAAATCTCCAGATGAGTG 276

QY 250 TATGCTTCAATTAACACCCCGATCGGACGCTAGCAATTTGGCTGTTCCTCAACTATGCA 309

Db 277 CGGCGCCAGCATAGCGCTTAAGCGGGAGCTACGGATCACCAACGCGCGGATGATGCG 336

QY 310 CAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAGATACCCAGTTTGAAGCTG 369

Db 337 GCGCAGGTATCATTTCCAGCGTGGCGGGTTCGCCGCTGCTCATCGCGCTGCGGGTG 396

QY 370 AATCTGTTTACCGGAATTCAGCCCGGACCTGATTCGCGAGCGTCTGGATGTGGTATC 429

Db 397 CGCCAGTCTCTCTCCATCATGCGGACCTCATTTTCAAGCAATTTGATGTGGGATA 456

QY 430 CGGCTGCGCGGCTTGCAGATTCAGCTGTTTTCGCGGCTGCGGCGGATGCAATG 489

Db 457 CGGCTGCTGCTGCGGATTCGCTATCGGCGACGCGGATGACCAAGTTTCGCGATC 516

QY 490 GTGCTGCGCGCGGAAAGTATCTCACAAATAGGATACCGGAAACCCGCG--- 545

Db 517 CTGCGCGTCTCTCGGCTGGCTGGAGCCATCCGCTGAGACGCTGCTCGATCTG 576

QY 546 ---CGATTGAGTAGTCAATCAATGCTTGAATACAGCGTGGCGCGGCAATGAATTTGA 602

Db 577 GCGCAGCTGAGTGGATTAATTCAGAGCGCTGCGGACGCGCTGCGCTGCGAGCTGCGC 636

QY 603 ACTGATCGCACCGAGGATCTGACTCGCTGATCCCAAGGAATTTGTGACTAA 662

Db 637 ACGATCACAGACAGGTGATTCGCGATAGCCAGCGCGCGGTTTCTG-CTGA 695

QY 663 TGATCCGATGACGCTGGTGGCTGAGCGGGTGGCGGATCGCTACGTGCGGCT 722

Db 696 TAGCGCTACGCGCTGATGAGCTTTGCTCTGGCGGATGCGGCTGCTGCTGCGCGC 755

QY 723 GATGTGGGTATCAACAGATCAATCGTGGGAGCTGGAGATCTTCTGCGCGGTTACCA 782

Db 756 ATGGCTGGTAGCAAAAGGTCGCGAGCGGGAGCTGGTGGCGCTCTTCCCGGAATACCA 815

QY 783 GTCAGATCCACGCGCGGTTTATCGGTTATATACCGAAAGATTAAGCTGCCCTGAAAGT 842

Db 816 TTTCCACACAGCGGGGTGATGCCCTTTATCCGACTCGAGCATTTTCCCGACCGCGGT 875

QY 843 ACGGTGCTGATCACTGCTGACGATTTATTTGTTGA 881

Db 876 GCGGGCATTTTATGACTTCTCGCGGAGAGGTTCGGTTA 914

RESULT 15

ABD02170/c

ID ABD02170 standard; DNA; 912 BP.

XX AC

XX AC

XX AC

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polynucleotide #774.

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

OS US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX P-PSDB; ABO68599.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 774; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABD01397-ABD17967 represent P. aeruginosa polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX Sequence 912 BP; 164 A; 256 C; 307 G; 145 T; 0 U; 0 Other;

Query Match 11.6%; Score 108; DB 11; Length 912;

Best Local Similarity 47.3%; Pred. No. 1.6e-22;

Matches 361; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 1 ATGGAACGACGCTAAACGACATGTCGGTGTTCGCAAGTAGTTCGATTTGGCTCTTTACC 60

Db 768 ATGATTTCGTTGAAGCATGGCGATCTTCGCCACTGTGTGTGACAGGGCTCGATGGCA 709

QY 61 GCCGCCGCCACAGACAGTACAGATGAGGTTTGTTCATCATAGTCAGACGGTATCAAACTG 120

Db 708 GCCCGCGCCAGAGCGCTGGCGCATGACCCCTTCGCGAGTCAGCCAGCAGATCCGCAAGCTG 649

QY 121 GAAGATGAGTTGACAGTAAAGCTGTTAAACCGTAGCACACGAGATTGGCTTGACCGAA 180

Db 648 GAGAGCGCGCCAGGTCACCTTTGCTGCATCGCACACCCCGCGGCTGACCTTACCGAG 589

QY 181 GCCCGTGAATTTACTACCGAGGCTGCCGTTCGTATGCTTCATGAGTCAGGATGTTTAT 240

Db 588 GCCCGCGAGGCTTCTATCGCAGTTGCGCGCAGATGCTGGCGATCGCCGAGGAGCGGAG 529

QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGAGCGCTAGCATTTGGCTTTCTTCA 300

Db 528 CGGCGGCTTGGCGAATGCGCGGATGCGCGGTTGGGTAGTTGCGCTTGGCGGCGCGGCTG 469

QY 301 ACTATGCAACAAATGTTCTCGCGGGCTTGACAGCCAAATGCTGAAAGATATCCAGGT 360

Db 468 GGGTTCTCGCGACGCTGATCACCCAGGCGCTGAAACCGTTGCTGGAGAACCCAGCGCAG 409

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 15:34:11 ; Search time 104.036 Seconds
(without alignments)
6353.876 Million cell updates/sec

Title: US-10-759-889-1

Perfect score: 930

Sequence: 1 atggacactaaacgcatt.....ggcgggaagaagaagtaa 930

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	621.4	66.8	933	4	US-09-489-039A-6503
2	122.4	13.1	948	4	US-09-252-991A-643
3	121.8	13.1	1026	4	US-09-252-991A-5186
4	112.6	12.1	927	4	US-09-489-039A-2078
5	110.4	11.9	957	4	US-09-252-991A-3952
6	110.4	11.9	1047	4	US-09-252-991A-4022
7	110.4	11.9	1089	4	US-09-252-991A-3893
8	108.6	11.7	915	4	US-09-489-039A-2786
9	108	11.6	912	4	US-09-252-991A-774
10	107.6	11.6	1017	4	US-09-252-991A-11975
11	107.6	11.6	1104	4	US-09-252-991A-11906
12	106.4	11.4	1029	4	US-09-252-991A-13686
13	106.4	11.4	1071	4	US-09-252-991A-13459
14	106.4	11.4	1809	4	US-09-252-991A-13887
15	103.2	11.1	933	4	US-09-489-039A-454
16	102.8	11.1	777	4	US-09-252-991A-5250
17	102.8	11.0	591	4	US-09-134-000C-11
18	101.4	10.9	915	4	US-09-252-991A-14968
19	101.4	10.9	1011	4	US-09-252-991A-14782
20	100.4	10.8	1062	4	US-09-252-991A-9086
21	100.2	10.8	1008	4	US-09-489-039A-3346
22	98.6	10.6	792	4	US-09-252-991A-12012
23	95.4	10.3	906	4	US-09-252-991A-8825
24	95.2	10.2	966	4	US-09-252-991A-3656
25	95.2	10.2	1017	4	US-09-489-039A-1578
26	95.2	10.2	1026	4	US-09-252-991A-3611
27	95.2	10.2	1350	4	US-09-252-991A-3851

28	92.6	10.0	1029	4	US-09-252-991A-153	Sequence 153, App
29	92.6	10.0	1110	4	US-09-252-991A-173	Sequence 173, App
30	92.6	10.0	1251	4	US-09-252-991A-163	Sequence 163, App
31	92.6	10.0	1611	4	US-09-252-991A-183	Sequence 183, App
32	91.2	9.8	774	4	US-09-252-991A-14788	Sequence 14788, A
33	91.2	9.8	957	4	US-09-252-991A-14537	Sequence 14537, A
34	91.2	9.8	1221	4	US-09-252-991A-14829	Sequence 14829, A
35	90	9.7	711	4	US-09-252-991A-12330	Sequence 12330, A
36	90	9.7	930	4	US-09-252-991A-12091	Sequence 12091, A
37	90	9.7	948	4	US-09-252-991A-12320	Sequence 12320, A
38	89.2	9.6	882	4	US-09-252-991A-8694	Sequence 8694, App
39	89	9.6	1008	4	US-09-252-991A-9548	Sequence 9548, App
40	89	9.6	2082	4	US-09-252-991A-9500	Sequence 9500, App
41	87.2	9.4	933	4	US-09-252-991A-9642	Sequence 9642, App
42	87.2	9.4	1095	4	US-09-252-991A-9724	Sequence 9724, App
43	87.2	9.4	1614	4	US-09-252-991A-9523	Sequence 9523, App
44	83.2	8.9	459	4	US-09-252-991A-14834	Sequence 14834, A
45	79.8	8.6	969	4	US-09-252-991A-345	Sequence 345, App

ALIGNMENTS

RESULT 1
US-09-489-039A-6503
; Sequence 6503, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Brelon et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004.001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6503
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-6503

Query Match	66.8%	Score 621.4;	DB 4;	Length 933;
Best Local Similarity	79.8%	Pred. No. 1.7e-190;		
Matches 733;	Conservative	0;	Mismatches 186;	Indels 0;
Gaps	0;			
Qy	1	ATGGAACGACTAAACCCATGTCGGTGTTCGCCAAAGTAGTTGAATTTGGCTTTTACC	60	
Db	4	ATGGAACGATTAAACCGCATGTCGGTTTCGCCAAAGTGTGCGATTAGCTTCACTCACC	63	
Qy	61	GCGCCGCCAGACGAGCTACAGATGAGCGTTTCCTCCATCAGTCAGACGGTATCAAACTG	120	
Db	64	GCAGCGCCCGCAGCTTCAGATGAGCGTTTCCTCCATCAGTCAGACGGTATCAAACTG	123	
Qy	121	GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCACAGCAGCATGTCCTGACCGAA	180	
Db	124	GAAGATGAGTTGCGAGTAAAGCTGTCACCGCAGCAGCCGCGAGCTTGGCTTAACGTAA	183	
Qy	181	GCGGTAGATTTACTACCGAGGCTGCGTCGTATGCTTCATCAAGTGGAGGATGTTTCA	240	
Db	184	GCAGTAAATCTATTACCGAGGCTGCGTAGGATGCTGTTTGAAGCGAGGATGTCAT	243	
Qy	241	GAGCACTGTATGCTTTCAATAACACCCCATCGGACGCTAGCGCATTTGCTGCTTTTCA	300	
Db	244	GAGCAATTTACGATTTAAACATACCCCATCGGACGCTGCGCATCGGCTCTCTTCA	303	
Qy	301	ACTATGCCACAAATGTTCTCGCGGGCTGACGCCAAATGCTGAAAGAAATCCAGGT	360	
Db	304	ACTATGCCACAAATGTTCTCGCGGGCTGACGCCAAATGCTGAAAGAAATCCAGGT	363	
Qy	361	TTGAGCGTCAATCTGGTTACCGGAAATTCACGCCCGGACCTGATTCGCGACGCTCTGAT	420	

364	Db	CTCTCTGTTAAATCTGGTCAACGGATACCTGCGCCAGACTGATCGCCACCGGCGCTTGAC	423
421	QY	GTGGTGATCCCGGTCGGCGCGCTTCAGGATTCAGCCGTGTTTCCCGCGCGTCTGGCGCG	480
424	Db	GTGGTGATCCCGGTCGGCGCGCTTCAGGATTCAGCCGTGTTTCCCGCGCGTCTGGCGTCA	483
481	QY	ATGCCAATGGTGGTGTGGCGCGCGGAAAGCTATCTCACAAATACGGCATACGGGAAAA	540
484	Db	ATGCCAATGGTATATATGGCGCGCTAAAAGCTATCTGGCGAGCTGGCAATCCAGAGAA	543
541	QY	CCGCGCGAATTTGAGTAGTCAATTCATGGCTTGAATACAGCGTGGCGGCCCGACAAATGAATTT	600
544	Db	CCGCGCGACCTTGCCGGACATGCTGGCTGGAATATAGCTGCGCGCGGACAAATGAGTTT	603
601	QY	GAACTGATCGCACCGGAAGGATCTGACTCGCCTGATCCCAAAAGGAGATTGTGACT	660
604	Db	GTCATATTCGCGCGGAAGGATCTGACCCGCTGACCCCGCAGGGGCGCTTCGTGACC	663
661	QY	AATGATCCGATGACGCTCGTGGCTGACGCGCGGTGCGCGGATCGCTACGTGCG	720
664	Db	AACGACCTATGACGCTGGTGGCTGACGCGCGGGTGGGTATCGCTATGTACCC	723
721	QY	CTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCGCGGTAC	780
724	Db	TTAATGTGGGCGATCGAAGAGATCAACCGCGGTGACTTGAGATCCTGCTGCCAGCTAT	783
781	QY	CAGTCAGATCCACGCCCGGTTTATGCGTTATATACCGAAAAAGATAAGTCGCGCTGAAG	840
784	Db	CAGTCGATCCGCGACCGGTCTACGCCCTGTATACGGAAAGGATAAACTGCGCGCTGAA	843
841	QY	GTCAGGTGCTGATCAACTCGCTGACGGATTTTGTGCTGAGTCGCTGAATTTGTTTCAG	900
844	Db	GTCAGGTCTGTATCAACTATCTGACGGAGTATTTGTGACGTGGCGAAGATCTATCAG	903
901	QY	GAGATGCACGGCGCGGA	919
904	Db	GGCATGCACGGCGCGGA	922

RESULT 2

```

RESOLI 2
US-09-252-991A-643
; Sequence 643, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 643
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-643

```

	Query Match	13.1%;	Score 122;	DB 4;	Length 948;
	Best Local Similarity	46.9%;	Pred. No. 4e-29;		
	Matches 417;	Conservative	0;	Mismatches 470;	Indels 3; Gaps 1
QY	1	ATGGAACGACTAAACGCGATCTCGTGTCTTGGCCAAAGTAGTTGAAATTTGGCTCTTTTACC	60		
Db	40	ATGGATTTCGCTGAAAGGATGCGCATCTTCGCCACTGTGTGGACAAAGGCTCGATGGCA	99		
QY	61	GGCGCGCCGACAGACGCTACAGATGAGCGTTTTCGTCCTACGTACAGACGCGTATCAAAACGT	120		
Db	100	GGCGGGCCCGACAGCTGGGATGACCCCTTCGCGAGTCAGCCAGACAGATCCGCAAGCTG	159		

121	GAAGATGAGTTGCGAGGTAAAGCTGTTTAAACCGTAGACACACGACGATTTGGCCTGACCGAA	180
160	GAGAGCCGCGCCAGGTCACTTTCTGTCATCGACACACCGCGGCTGACCTTGACCCGAG	219
181	GCCGGTAGAATTTACTACCAAGGCTGCGTGTATGCTTCATGAAGTCAGGATGTTAT	240
220	GCCGCGGAGGGTTCTATTCGAGTTTGGCGCAGATGCTGGGATCGCCGAGGAAGCCGAG	279
241	GAGCAATGTTATGCTTCAATAACACCCCATCGGACGCTACGCAATCGCTGTTCTTCA	300
280	CGCGGCTTGCGGAATGCGCGCATCGCCGGTGGGTGAGTTGCGCTGCGCGCGCGGTG	339
301	ACTATGGCACAAATGTTCTCGCGGGCTCACAGCCAAAATGCTGAAGAAATACCAAGT	360
340	GGGTTCCTCCGGCAGCTGATCACACGAGCGCTGAACCGTTGCTGGAGAAACACACGCCAG	399
361	TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCGCCGAACTGATTCGCCAGCGTCTGAT	420
400	TTGCGCTGCGAGCTGTTCTTCAGGACGAGCGCATCGATCTGCTGCCGAGCGCATCGAC	459
421	GTGGTGATCCCGGTCGGCGGTTGCAGGATTCAGAGCTGTTTCCCGCGCTCTGGGCGCG	480
460	CTGGCGATACGCGTCGGCAATCTCGCGCATCTCAGAGCTGGTGGCGGCGCACCTCGGCGAT	519
481	ATGCCAATGCTGGTGGCGCGGCAAAAGCTATCTCACACAATACGGCATACCGGAAAAA	540
520	TGGAGCAGCTGCTCTGGCGGCGCGGCTATTTGGCCAAAGGGGCGCCATCAACGCT	579
541	CCGCGCGAATTTGAGTAGTATCATAGGCTTT---GAAATACAGGTGGCGGCCGACAAATGAA	597
580	CGGAGCAGTTTGAACGAGGTGCGACTGTGATTTCCCTGAACACCTCGAACACCTCAACAC	639
598	TTTGAACGTATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGAGAAAGATTGTG	657
640	CTGACGCTACGCGGCGCGGGGGAAGTCTGCAAACTGCGCTCGAGCCACGGGTGGCC	699
658	ACTAATGATCGATGACGCTGGTGGCTGAGCGGGGGTCCGGGATCGCTACGCTG	717
700	GCCAACGGGATGCTCGCGGTGCGTCACTTCAACCTCGAAGGACTCGGCTGCTACCAG	759
718	CGGCTGATGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCGCGGT	777
760	CGGCTGCGGAGGTGCGGACGCGCTGAACGACGGGCGCTGCAAGCTATTGCGGAG	819
778	TACCACTCAGATCCACGCCCGGTTTATGCTTATATACGAAAAAGATAAGCTGCGGCTG	837
820	TGAGGATTCGCGGCTCGGCATCTATGCGTGAACGCCAGCGCGGAGCCAGCGCGCC	879
838	AAGGTACAGTCTGATCAACTCGCTACCGGATTTTGTGAGTCTCG	887
880	AAGGTGAAGTCCGATCGAGCGCTCGCGCGAGCGTTCGCCACGAGCACGG	929

RESULT 3
 US-09-252-991A-5186
 ; Sequence 5186, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107156.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5186
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa

QY 841 GTACAGTCTGATCAACTCGCT 863
Db 856 CTGACGCGTATTATCGCCACCT 878

RESULT 5

US-09-252-991A-3952
; Sequence 3952, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3952
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3952

Query Match 11.9%; Score 110.4; DB 4; Length 957;
Best Local Similarity 50.0%; Pred. No. 2.3e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
QY 1 ATGGAACGACTAAACCGCATGTCGGTGTGTTGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 34 ATGACCGGTTTCGATCGGATGACGGCTTTGCCGCTGTGTGATACCGGAGCTTACC 93
QY 61 GCGCGCCGACAGCAGTACAGATGAGCGTTTGTCTCATGATCAGTACAGCGGTATCAAACTG 120
Db 94 AAGCGCGCGAGACGCTGCACATGACGCGCACCGCATCCAGTTCACCGTGTGTCAGCAACTG 153
QY 121 GAAGATGAGTTGACAGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTTGACCGAA 180
Db 154 GAAGCGCGCTGCGGTTACGCTGTCTCAACCGCACACCGGCAAGCTCAACTGACCGCC 213
QY 181 GCGGTAGAAATTTATACAGGCGCTGCGTGTATGCTTTCATGAAAGTGAGGATTTTCAAT 240
Db 214 GACGCGCGCGCTTACTACAGCGGGTGTGATCGTCTGCTGGCGGAACTGGACGATGCCGAG 273
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTCTTCA 300
Db 274 ACAGCGCTCTCAGCGCGCGCTTGGCGCGCGCGCTGCGGCGCTTCCGATCGACTACCGAGC 333
QY 301 ACTATGGCACAAATTTCTCGCGCGGCTGACAGCCAAATCTGAAAGATACCCAGGT 360
Db 334 CGCTGGCGCGCTGCTGTATCCCGGCTTGGCGGCTTCCATGCGCGCTATCCGGAG 393
QY 361 TTGAGCGTCAATTTGTTACCGAATTCAGCCCGGCTGATTCGCGACGCTTGGAT 420
Db 394 ATCCAGATCGAATTTGGCGGCTGATCGGCTGCTGATCGGCTGATCGAGAGAACGTGAC 453
QY 421 GTGGTATCGCGCTCGGCGGCTTGCAGGATTCAGGCTGTTTCCCGCGCTGCGCGCG 480
Db 454 TGGCTCATCGCGCGCGGCTGATCGCGCGGCTGATCGCGCGGCTGCGCGAG 513
QY 481 ATGCCAATGTTGTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAATA 540
Db 514 CTGCAACTGGGGTCTACGCGGACCGGCTTACTTTCAGCGCGCGCGGCTGCGCGAGCCAC 573
QY 541 CCGCGCGATTG 552
Db 574 CCGCGGAACTG 585

RESULT 7

US-09-252-991A-3893/c
; Sequence 3893, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

RESULT 6

US-09-252-991A-4022
; Sequence 4022, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4022
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4022

Query Match 11.9%; Score 110.4; DB 4; Length 1047;
Best Local Similarity 50.0%; Pred. No. 2.4e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
QY 1 ATGGAACGACTAAACCGCATGTCGGTGTGTTGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 44 ATGACCGGTTTCGATCGGATGACGGCTTTGCCGCTGTGTGATACCGGAGCTTACC 103
QY 61 GCGCGCCGACAGCAGTACAGATGAGCGTTTGTCTCATGATCAGTACAGCGGTATCAAACTG 120
Db 104 AAGCGCGCGAGACGCTGCACATGACGCGCACCGCATCCAGTTCACCGTGTGTCAGCAACTG 163
QY 121 GAAGATGAGTTGACAGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTTGACCGAA 180
Db 164 GAAGCGCGCTGCGGTTACGCTGTCTCAACCGCACACCGGCAAGCTCAACTGACCGCC 223
QY 181 GCGGTAGAAATTTATACAGGCGCTGCGTGTATGCTTTCATGAAAGTGAGGATTTTCAAT 240
Db 224 GACGCGCGCGCTTACTACAGCGGGTGTGATCGTCTGCTGGCGGAACTGGACGATGCCGAG 283
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTCTTCA 300
Db 284 ACNAGCTCTCAGCGCGCGCTTGGCGCGCGCGCTTCCGATCGACTACCGAGC 343
QY 301 ACTATGGCACAAATTTCTCGCGCGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360
Db 344 CCGCTGGCGCGCTGCTGTATCCCGGCTTCCGCGAGTTCATGCGGCTATCCGAG 403
QY 361 TTGAGCGTCAATTTGTTACCGAATTCAGCCCGGCTGATTCGCGACGCTTGGAT 420
Db 404 ATCCAGATCGAATTTGGCGGCTGAGCGGATCGGCTGCTGATCGCTGATCGAGAGACGTGAC 463
QY 421 GTGGTATCGCGCTCGGCGGCTTGCAGGATTCAGGCTGTTTCCCGCGCTTGGCGCG 480
Db 464 TGGCTCATCGCGCGCGGCTGATCGCGAGCTGCGGAGTTCGATAGCGCGGCTGCGCGAG 523
QY 481 ATGCCAATGTTGTGTCGCGCGGAAAGCTATCTCACAAATACCGCATACCGGAAATA 540
Db 524 CTGCAACTGGGGTCTACGCGGACCGGCTTACTTTCAGCGCGCGCGGCTGCGCGAGCCAC 583
QY 541 CCGCGCGATTG 552
Db 584 CCGCGGAACTG 595

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252.991A
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 3893
;; LENGTH: 1089
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3893

Query Match 11.9%; Score 110.4; DB 4; Length 1089;
Best Local Similarity 50.0%; Pred. No. 2.4e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGACATGTCGGTGTGTCACAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db |||||
QY 61 GCCGCGCCAGACAGCTACAGATGAGCGTTTTCGTCATCAGTCAGACGGTATCAAAATG 120
Db |||||
QY 921 AAGGCGCGGAGACGCTGCACATGAGCGCAGCCAGCGTCAACCAGTTGGTCCAGCACTG 862
Db |||||
QY 121 GAAGATGAGTTGAGCTGAAGCTTTAAACCGTAGCACAGCGAGCTATGCGCTGACCGAA 180
Db |||||
QY 861 GAAGCGGGCTGCGGGTACGCGTCTCAACCGCACCGCGCAAGCTCAACCTGACCGCC 802
Db |||||
QY 181 GCCGGTAGAATTTACTACGAGGCTGCCGTGCTATGCTTCATGAAGTCAGAGATGTTTAT 240
Db |||||
QY 801 GACGGCGCCCTACTACGAGCGGGTGGTAGCTCTCTGGCGGAATCGACGATCGCGAG 742
Db |||||
QY 241 GAGCAACTGTATGCCCTTAATAACACCGCATCGGACGCTAGCAATGGCTGTTCTTCA 300
Db |||||
QY 741 ACCAGCTCTCCAGCGCCGCCCTTGGCGCGCCCGCGCGCTCTCGGATCGACGTACCGAGC 682
Db |||||
QY 301 ACTATGGCAAAATGTTCTGCGCGGGCTGACAGCCAAATGCTGAAGAATACCCAGGT 360
Db |||||
QY 681 CGCTGCGCGGCTGCTGTGATCCCGGGGTGTCGCGATTCATGCGCGGCTATCGGAG 622
Db |||||
QY 361 TTGAGCGTCAATGTGTTACCGAATTCAGCGCCCGACCTGATTCGCGACGGTCTGGAT 420
Db |||||
QY 621 ATCCAGATCGACCTGGCGGTGACGATCGGCTCTGACCTGATCGACGAGAAGCTCGAC 562
Db |||||
QY 421 GTGGTGTATCGGCTGCGCGGTGTCAGGATTCAGGCTGTTTCCGCGCTGCGGCGG 480
Db |||||
QY 561 TGGCTCATCGCGCGCGGAACTCGCGACCGTGTGATAGCGCGCGGCTCGGCGAC 502
Db |||||
QY 481 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
Db |||||
QY 501 CTGCAACTGGGGTCTACGCGGACCGGCTACTTGCAGCGCGCGGCTGCGGAGCCAC 442
Db |||||
QY 541 CCGCGGATTTG 552
Db |||||
QY 441 CCGCGGAACTG 430
Db |||||

RESULT 8

US-09-489-039A-2786
; Sequence 2786, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 2786
;; LENGTH: 915
;; TYPE: DNA
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2786

Query Match 11.7%; Score 108.6; DB 4; Length 915;
Best Local Similarity 46.9%; Pred. No. 8.5e-25;
Matches 412; Conservative 0; Mismatches 459; Indels 8; Gaps 2;

QY 10 CTAAACGCAATGTCGGTGTGTCGCAAAAGTAGTTGAATTTGGCTCTTTTACC CGCGCGCC 69
Db |||||
QY 37 CTTGAGCGCTGGCCATGTTTATCGCGTGTGATGCGCGCAGTTTACCTCGCGCGCA 96
Db |||||
QY 70 AGACAGCTACAGATGAGCGTTTTCGTCATCAGTCAGACGGTATCAAACTGGAAGATGAG 129
Db |||||
QY 97 GCGGCGCTGGGCGAGACCAAGCGGTGCTCAGCTTTAACGTACGCGAGCTGGAATGAG 156
Db |||||
QY 130 TTGCAAGTAAAGCTGTTAAACCGTAGCACAGCGAGCTATGCGCTGACGGAAGCGGTGAG 189
Db |||||
QY 157 CTGGGGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
Db |||||
QY 190 ATTTACTACAGGCTGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249
Db |||||
QY 217 CTGTTCTATCAGCGCGGGTGGCGCTCTCTGAACCGCGGGAATCTCCAGGATGAAGTG 276
Db |||||
QY 250 TATGCGTTCAATACACCCCGCATCGGACGCTACGCAATGGCTGCTTTTCAACTATGCGCA 309
Db |||||
QY 277 CGGCGCAACCATAGCGGCTTAAAGCGGGAGCTACGGATCACCAACGCGCGAGTATGCG 336
Db |||||
QY 310 CAAATGTTCTCGCGGCTGACAGCAAAATGCTGAAGAATACCCAGGTTTCAGCGTC 369
Db |||||
QY 337 GCGCAGTTATCATTCAGCGCTGCGCGCTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 396
Db |||||
QY 370 AATCTGTTACCGAATTCAGCGCCCGACCTGATTCGCGCGGTCTGGATGCTGGTATC 429
Db |||||
QY 397 CGCCACGTTCTCTCTCCCATCATCGCGACCTCATTTTCAAGAACGATTTGATGTCGCGATA 456
Db |||||
QY 430 CGCGTCCGCGGTTGAGGATTCAGCGCTGTTTCCCGCGCTGCGCGGCGATGCCAATG 489
Db |||||
QY 457 CGCTCGGTGCTGCGCGGATTTCCCGCTATCGGCGGAGCGGATAGCGAGTTTCGCGATC 516
Db |||||
QY 490 GTGGTGTGCGCGGCAAAAGCTATCTCACAAATACCGGCATACCGGAAAAACCGC --- 545
Db |||||
QY 517 CTGCGCGTGGCTCTCGCGCTGGCTGCAAGCCATCCGGTGCAGACGCTGCTCCGATCTG 576
Db |||||
QY 546 ---CGATTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
Db |||||
QY 577 GCGCAAGCTGAGTGGATTTATTCAGAGCGCTGCGGACGCGCGCTGCGCTGCGCAGCTCGCG 636
Db |||||
QY 603 ACTGATCGCACCGGAAAGGATCTCGACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
Db |||||
QY 637 ACAGATCACAGACAGAGTGGATTTCCGATAGCCAGCGCGCGCGGTTTCTG-CTGA 695
Db |||||
QY 663 TGATCCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
Db |||||
QY 696 TAGCGCTACGCGCTGATGAGCTTTGCTCTGCGGGGATGCGGGCTGCTGCTGCTGCTGCTGCTG 755
Db |||||
QY 723 GATGCGGTGATCAACGAGATCAATCTGCGGGAGCTGAGATCTGCTGCTGCGGTTACCA 782
Db |||||
QY 756 ATGCTGTTAGCCAAAAGGTGGCGAGCGGGAGCTGCTGCGGCTCTCTGCGGAATACCA 815
Db |||||
QY 793 GTCAGATCCAGCGCGGTTTATGCGTTTATACCGAAAAAGATAAGCTGCGCTGGAAGGT 842
Db |||||
QY 816 TTTCCCAACAGCAGCGCGGTGATGCCCTTTATCCGACTCGCAGCATTTGCGGAGCCCGGT 875
Db |||||
QY 843 ACAGTGTGATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
Db |||||
QY 876 GCGGGCATTTATTGACTTCTGCGCGGAGAGGTGCGTTA 914
Db |||||

RESULT 9

US-09-252-991A-774/c

; Sequence 774, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 774

; LENGTH: 912

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-774

Query Match 11.6%; Score 108; DB 4; Length 912;

Best Local Similarity 47.3%; Pred. No. 1.3e-24;

Matches 361; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

Qy 1 ATGGAACGACTAAAGCGCATGTCGGTGTGTTGGCAAGTAGTTGAATTTGGCTCTTTTACC 60

Db ATGGAATTCGCTGAAGGCATGGCGATCTTCGCCACTGTGTGACAAAGGCTCGATGCA 709

Qy 61 GCGCGCCGACAGACGTACAGATGAGCGTTTGTGTCATCAGTCAGACGATCAAACTG 120

Db GCGCGCGCCAGACGCTGGGCGATGACCTCTGGCAGTCAGCCAGCAGATCCGCAAGCTG 649

Qy 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACAGCAGCATTTGGCTTGACCGAA 180

Db GAGAGCGCGCCAGCTGACTTTGTGTCATCGCACACCCCGCGGTGACCTGACCGAG 589

Qy 181 GCGGTAGAAATTTACTACAGGCGTCCGTCGTATGTTTCATGAAGTCAGGATGTTAT 240

Db GCGCGCGAGCGGTCTCTATCGAGTTGCGCGCAGATGCTGGCGATCGCCGAGGAAGCCGAG 529

Qy 241 GAGCAACTGTATGCTTCAATACACCCCATCGGACGCTACGCAATTTGGCTGTTCTTCA 300

Db CGCGCGCTTGGCGAATGGCGCGATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 469

Qy 301 ACTATGGCACAAATGTTCTCGCGCGGCTGACAGCCAAATGCTGAAGAAATACCCAGGT 360

Db GGGTCTTCGCGCACGCTGATCAACAGGCGCTGAACCCGTGCTGGAGAACACCCGCCAG 409

Qy 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGACCTGATTCGCGACCGTCTGGAT 420

Db TTGCGCTGCGAGTGTCTTCCAGGACGAGCGCATCGATCTGTCGCCGAGCGCATCGAC 349

Qy 421 GTGGTATCGCGCTCGCGCGGTGTCAGGATTCAGGCTGTTTTCGCCGCGTCTGGCGCG 480

Db CTGGCGATAGCGTGGGCAATCTCGCGACTCGAGCTGTTGGCGCGCCACCTCGCGCAT 289

Qy 481 ATGCAATGTTGTGTGCGCGCGGAAAGCTATCTCACACATACGCGCATACCGGAA 540

Db TGGAGCAGCGTGTCTGCGCGCGCGCGGCTATTTGCGCAAGCGGCGCGCCATCAACCGT 229

Qy 541 CCCGCCGATTTAGTAGTCAATTCATGCTTTGAAT ---ACAGGTGGCGCCGACATGAA 597

Db CCGGAGCAGTTGACCGAGTGTGATGATTTCCCTGAAACACCTGACACCTCAACAC 169

Qy 598 TTTTAACTGATCGACCGGAAGGATCTCGACTGCGCTGATCCCAAGGAAGATTTGTG 657

Db CTGACGCTACGCGCGCGCGCGGGAAGTCTGCAAACTGCGGCTGAGGACCGGCTGGCC 109

Qy 658 ACTAATGATCCGATGACGCGGTGGTGGCTGGCTGACGCGCGGTGCGGGATCGCTACG 717

Db GCCAACGGGATGCTGCGGCTGGCTGAGTTCAACCTCGACGAGTCTGCGGTGCTTACCG 49

Qy 718 CCGCTGATGGGTGATCAACGAGATCAATCGTGGGAGCTGCA 761

Db 48 CCGCTCCGGAGGTGCGGACGCGCTGACGACGGCGCTGCA 5

RESULT 10

US-09-252-991A-11975

; Sequence 11975, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11975

; LENGTH: 1017

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11975

Query Match 11.6%; Score 107.6; DB 4; Length 1017;

Best Local Similarity 46.1%; Pred. No. 1.9e-24;

Matches 399; Conservative 0; Mismatches 464; Indels 3; Gaps 1;

Qy 2 TGGAACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACCG 61

Db TGGACACATTCGAAGGCATGCGGTGTTCGCCAGGTGTTGGACACGCGGCGCTTCACTT 122

Qy 62 CCGCGCGACAGACGCTACAGATGAGGTTCGTCCTCATGTCAGACGGTATCAAACTGG 121

Db 123 CCGCGCGACAGGCGCTGACCTGTCCACGCGCAGGTATCGCGGCTGATCTCGACTGG 182

Qy 122 AAGATGAGTTGCAAGTAAAGCTGTTAAACCGTAGACACGCGACATTCGCGCTGACGAG 181

Db 183 AGGCGCACCTGACAGACGCGCTGTCACGCGCACCCCGCGCTGGCGCTGACCGAG 242

Qy 182 CCGGTAGAAATTTACTACAGGCTGCGCTGATGTTTCAATGAAGTGCAGGATGTTTCATG 241

Db 243 CCGCGCAGCGTTTACTTGAACCGTTCGCGCGGATCTTCGAGGACATCGAGTGCAGG 302

Qy 242 AGCAACTGTATGCTTCAATTAACACCCCATTCGCGAGCGTACGCAATGGCTGTTTCAA 301

Db 303 CCGAAGCGCGCGCGCACATCGGCCCTGCGCGCTTGCCTGCGCTGAGTCTGCTGATGG 362

Qy 302 CTATGGCACAAATGTTCTCGCGGCTGACACCCAAATGCTGAAGATATCCAGGTT 361

Db 363 GCATGGCCAGCATACCTGGTGGGATGATCGCGCTACCGGCGACTGTTCCCGAGC 422

Qy 362 TGACGCTCAATCTGTTTACCGGAATTCAGCCCGCCGACCTGATTCGCGAGCTTGGATG 421

Db 423 TGGTGTGAGCTGACCTGTCCAGCGCAACCGGACATGCTCGAGGAAGGCGAGGACG 482

Qy 422 TGGTGTGAGCTGCG---GGCGTTGAGGATTCAGCTGTTTCCGCGCTGCGGCG 478

Db 483 TGCTGATCATCCGCGAGCGCGAGTTGCCGACTCGGAGTTTCGTCGCCGCGCTGGGCA 542

Qy 479 CGATGCCAAATGTTGCTGCGCGCGGAAAGCTATCTCACCAATACCGCATACCGGAAA 538

Db 543 GCATTCACAGGCTGCTGCGCGAGTCCCGGCTACCTGCGAGCGCGGCGTCCAGCT 602

Qy 539 AACCCGCGAATTTGAGTAGTCAATTCATGCTTGAATACAGCTGCGGCGCGCAATGAAT 598

Db 603 CCGTGGACGACCTGGAGCGCACTGCTGCTGCGGCTGCGAGGACCGGCGCTACCGGAGG 662

Qy 599 TTGACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGA 658

Db 663 GCTGATCTTCGCGACGAGCGGCGAGCGACCGTCAGCCGCGAGAACCTTCATGG 722
QY 659 CTAATGATCCGATGACGCTGGTGGCTGACGCGGGTCCCGGATCGCTACGTGC 718
Db 723 TCAATGTCCCGAGGTCATGGCGCCAGCGCGCCAGCGACCTCGGCATCGCCCTGCTGC 782
QY 719 CGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGCTGCCGCTT 778
Db 783 CCAGCTACGTGGCGCGCGCGCTGCGCAACGCGGAGCTGTACGGGTGTGCCCGGCC 842
QY 779 ACCAGTCAGATCACGCCCGGTTTATGCTTTATATACGAAAGATAGCTGCCGCTGA 838
Db 843 ATGCCATGCAACGAGCGGACATATGCGCTGTATCCCTCGCGGGGCTTACCTGGAGCGGA 902
QY 839 AGGTACAGGTGCTGATCAACTCGCTG 864
Db 903 AGATCCGACCTGGGTGATCTGTTG 928

RESULT 11

US-09-252-991A-11906
; Sequence 11906, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11906
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11906

Query Match 11.6%; Score 107.6; DB 4; Length 1104;
Best Local Similarity 46.1%; Pred. No. 26-24; Indels 3; Gaps 1;
Matches 399; Conservative 0; Mismatches 464;
QY 2 TGGACGATCAAAACGATGCGGTGTTGCGAAAGTATGAAATTTGGCTCTTTTACCG 61
Db 167 TGGACACATTCGAAGCATGCGGTGTTGCGCAGGCTGTGACAGCGGAGCTTCACTT 226
QY 62 CCGCGCCAGACGATCAGATGAGGTTTCGTCATCAGTCAGACGCTATCAAACTGG 121
Db 227 CCGCGCCAGCGCCCTGGACCTGTCCACCGCGCAGGTATCGCGGCTGATCTCCGACCTGG 286
QY 122 AAGATGAGTTGCAAGTAAAGCTGTTAAACCGTAGCACGACGATTCGCTGACCGAAG 181
Db 287 AGCGCACCTCGACAGCGGCTGTGACCGGACACACCGCGGCTTGGCTTACCGAGG 346
QY 192 CCGGTAGAAATTTACTACAGGCGTGGCTGCTATGCTTATGAAAGTGTGAGATGTTATG 241
Db 347 CCGCGGAGCGTTTACCTGGAAACGCTGCGCGGATCTCGAGACATCAGAGTTCGCGGAGG 406
QY 242 AGCACTGATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTCGCTGTTCTTCAA 301
Db 407 CCGAGCCCGCGCGGCGACATCCGCGCTTGGCGGCTTGGCGTGTGAGTGTGATGG 466
QY 302 CTATGGCAAAATGTTCTCGCGGCGTGAACGCAAAATGCTGAAAGATACCCAGTT 361
Db 467 GCATGGCGCAGCATCACCTGGTGGCGATGATCGCGCGCTACGCGGCACTGTTCCCGGACG 526
QY 362 TGAGCGTCAATCTGTTACCGAAATTCAGCCCGGACCTGATGCCCGAGCTGCTGATG 421
Db 527 TGGTATGAGCTGACCTGTCCCGGCAACCCGGAATGCTCGAGGAAGCGGAGCG 586

QY 422 TGGTATCCCGGTGCG---GGCGTTGAGAGTTCAGACCTGTTTTCGCGCGTCTGGCG 478
Db 587 TGGTATACACCGGAGCGCCAGTTGCCACTCGGAGTTCGTGCGCCAGCGCTGGGCA 646
QY 479 CGATGCCAAATGGTGTGGCGCGCGAAAGACTATCTCACAAATACGGCATACCGGAAA 538
Db 647 GCATCCACAGCGTCTCTGCGCAGTCCCGGCTACCTGCAGCAGCGCGGTGCCACGCT 706
QY 539 AACCCGCGGATTTGAGTAGTCAATTCATGGCTTGAATACAGCGTGGCGCGCGCAATGAAT 598
Db 707 CGGTGAGCAAGCTTGGAGCGCCACGTCTGCTGCGCTGCGAGTACCGGCTACCCGAGG 766
QY 599 TTGAATGATCGCACCGGAGGATCTCGACTCGCTCGCTGATCCCAAGAGAGATTTGTGA 658
Db 767 GCTGATCTTCGCGCAGCAGCGGCGAGCGCACCTTCAGCCCGCAGAACCTTCATGG 826
QY 659 CTAATGATCCGATGACCGCTGGTGGCTGTGACGCGGGGTGCGGGATCGCTACGCTGC 718
Db 827 TCAATGTGCGCGAGGTCTATGCGCCAGCGCGGCGCAAGCGCGACCTCGGCATCGCCCTGCTGC 886
QY 719 CGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTCTGCGCGGTT 778
Db 887 CCAGCTAGTGGCGCGCGCGCTGCGCAACGCGGAGCTGTACGGGTGTTGCCCGGCC 946
QY 779 ACCAGTCAGATCCACGCCCGCTTTATGCTTATATACGAAAGATAAAGTCCCGCTGA 838
Db 947 ATGCCATGCACCGGAGCCATCTATGCGCTGTATCCCTCGCGCGTTCCTGAGCGCA 1006
QY 839 AGGTACAGGTGCTGATCAACTCGCTG 864
Db 1007 AGATCCGACCTGGGTGATCTGTTG 1032

RESULT 12

US-09-252-991A-13686
; Sequence 13686, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13686
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13686

Query Match 11.4%; Score 106.4; DB 4; Length 1029;
Best Local Similarity 49.5%; Pred. No. 4.6e-24;
Matches 275; Conservative 0; Mismatches 281; Indels 0; Gaps 0;
QY 28 TTTGCAAAAGTAGTGAATTTGGCTCTTTTACCGCGCGCGCCAGACAGCTACAGATCAGC 87
Db 172 TTGTCAGTGTAGTCGACAGCGGTTCCATCAGCGCGCGCGGAAACAGATGGAGTACC 231
QY 88 GTTTCGTCATCAGTCAGCGGTATCAAACTGGAAGATGAGTTGCAAGTAAAGCTGTTA 147
Db 232 GCCTCGGCGCTCAGCGGAGCCCTCTCGGCTCTGAGCAAGCTCGGCAACCCCTGCTC 291
QY 148 AACCGTAGCACGCGAGCATTTGGCTTACCGAAGCGGTAGAATTTACTACCAAGGCTGC 207
Db 292 AACCGCACACCCCGCGGCGATGAGCTGACCGAGAGGCGGCTACTTCTTCGACCGGCA 351
QY 208 CGTCGTATGCTTCATGAAGTGCAGGATGTTTCATGAGCAACTGTATGCTTCAATAACAC 267

Db 352 CGGACAGATCTGAGAGATGAGGCGCTCGAGGAGCGCTGGCCCTCGCGCGGAGAGC 411
Qy 268 CCATCGGAGCGCTAGCGATGGCTGTTCTTCAACTATGGACAAATATGTTCTCGCGGG 327
Db 412 CCGCGCGGCGCTGGGATCAACGCGCTCGCCCTCATGCTGACGCGGTGTGCGG 471
Qy 328 CTGACAGCAGAAATGCTGAAAGATACCCAGGTTTGGAGGTCATCTGTTACCGGAAT 387
Db 472 CACATCGGCGAGTTCCGCGCGGTGTACCGGACATCCAGCTCGAACTCAACACCAACGAC 531
Qy 388 CAGCGCCCGACCTGATTCGCGACGCTCTGGATGTGGTATCGCGGTGGGCGGTTCGAG 447
Db 532 CTGATCATGACCTGTTGAGCAGCAGCAGCTGGGATCCGATCGGTGGCTCAGC 591
Qy 448 GATTCAGCGCTGTTTCCGCGCTCTGGGCGGATGCCAATGGTGGTGGCGCGGAAA 507
Db 592 GACTCCAGCTGACGCGGCGCGCTGGGCGGACGCGCTGCATCTCGCCAGCCCG 651
Qy 508 AGCTATCTCACAAATACGCGCATACCGGAAACCGCGGATTTGAGTAGTCAATTCATGG 567
Db 652 GAGTACTGCGCGCAGCGGACGCGCGGAGCATCGAGGGCTGCGAGGGGACACCCCTG 711
Qy 568 CTTGAATACAGCGTGC 583
Db 712 CTCGGCTTCAACCCAGC 727

RESULT 13

US-09-252-991A-13459/c
; Sequence 13459, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13459
; LENGTH: 1071

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13459

Query Match 11.4%; Score 106.4; DB 4; Length 1071;
Best Local Similarity 49.5%; Pred. No. 4.7e-24;
Matches 275; Conservative 0; Mismatches 281; Indels 0; Gaps 0;

Qy 28 TTGCGCAAGTAGTTGAATTTGGCTCTTTTACGCGCGCGCCGAGACAGTACAGATGAGC 87
Db 1035 TTGCTCAGTGTAGTCAGACCGGTTCCATCAGCGCGCGCGGAGAACAGATGGAGCTGACC 976
Qy 88 GTTTCGTCATCAGTCAGACGGTATCAAAATCGGAAGATGAGTTGACGTAAGCTGTTA 147
Db 975 GCCTCGGCGCTCAGCGGACCTCTCGCGCTCGAGCACAAGCTCGGACCAACCTGCTC 916
Qy 148 AACCTAGACACGAGAGATTGCTGACCGAGCGGTAGAAATTTACTACAGGGCTGC 207
Db 915 AACCGCACACCGCGGCGATGAGCTGACCGGAGGAGCGCGTACTTCTCTGACAGGCA 856
Qy 208 CGTCTGATGCTTCATGAAGTGCAGGATGTTATGAGCAACTGTATGCCTTCAATAACACC 267
Db 855 CGGCAGATCTGACGAGATGGAGCGCTCGAGGAACGCTTGGCTGCGCGCGGAGAGC 796
Qy 268 CCCATCGGAGCGGTAGCGATTTGCTGTTCTTCAACTATGGCAAAATGTTCTCGCGGG 327
Db 795 CCGCGCGGCGCTGCGGATCAACCGCGCTCGCCCTTCATGCTCAGCGGTGTGCGG 736

Qy 328 CTGACAGCAGAAATGCTGAAAGATACCCAGGTTTGGCGGTCAATCTGTTACCGGAAT 387
Db 735 CACATCGGCGAGTTCCGCGCGGTGTATCCCGGACATCCAGCTCGAACTCAACACCAACGAC 676
Qy 388 CAGCGCCCGACCTGATTCGCGACGCTCTGGATGTGGTATCGCGCTCGCGCGGTTCGAG 447
Db 675 CTGATCATGACCTGTTGGAGCAACGACCGAGTGGCATCGCATCGTGGCTCAGC 616
Qy 448 GATTCAGCGCTGTTTCCGCGCTCTGGGCGGATGCCAATGGTGGTGGCGCGGAAA 507
Db 615 GACTCCAGCTGACGCGCGCGCTGGGCGGACGCGCTGCATCTCGCCAGCCCG 556
Qy 508 AGCTATCTCACAAATACGCGCATACCGGAAACCGCGGATTTGAGTAGTCAATTCATGG 567
Db 555 GAGTACTGCGCGCAGCGGACGCGCGGAGCATCGAGGGCTGCGAGGGACACCCCTG 496
Qy 568 CTTGAATACAGCGTGC 583
Db 495 CTCGGCTTCAACCCAGC 480

RESULT 14

US-09-252-991A-13887
; Sequence 13887, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13887
; LENGTH: 1809

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13887

Query Match 11.4%; Score 106.4; DB 4; Length 1809;
Best Local Similarity 49.5%; Pred. No. 6.4e-24;
Matches 275; Conservative 0; Mismatches 281; Indels 0; Gaps 0;

Qy 28 TTGCGCAAGTAGTTGAATTTGGCTCTTTTACGCGCGCGCCGAGACAGTACAGATGAGC 87
Db 11 TTGCTCAGTGTAGTCAGACCGGTTTCATCAGCGCGCGCGGAGAACAGATGGAGCTGACC 70
Qy 88 GTTTCGTCATCAGTCAGACGGTATCAAAATCGGAAGATGAGTTGCAAGTAAAGCTGTTA 147
Db 71 GCCTCGGCGCTCAGCGGACCTCTCGCGCTCGAGCACAAGCTCGGACCAACCTGCTC 130
Qy 148 AACCTAGACACGAGCAGTATGCTGACCGAGCGGTAGAAATTTACTACAGGGCTGC 207
Db 131 AACCGCACACCGCGGCGATGAGCTGACCGGAGGAGCGGCTACTTCTTCGACAGGCA 190
Qy 208 CGTCTGATGCTTCATGAAGTGCAGGATGTTTCATGAGCAACTGTATGCTTCAATAACACC 267
Db 191 CGGCAGATCTGACGAGATGGAGCGCTCGAGGAACGCTGCGCCCTGCGCGCGGAGAGC 250
Qy 268 CCCATCGGAGCGGTAGCGATTTGCTGTTCTTCAACTATGGCAAAATGTTCTCGCGGG 327
Db 251 CCGCGCGGCGCTCGGATCAACCGCGCTCGCCCTTCATGCTCAGCGGTGTGCGG 310
Qy 328 CTGACAGCAGAAATGCTGAAAGATACCCAGGTTTGGAGGTCAATCTGTTACCGGAAT 387
Db 311 CACATCGGCGAGTTCCGCGCGGTGTACCGGACATCCAGCTCGAACTCAACACCAACGAC 370
Qy 388 CAGCGCCCGACCTGANTGCGGACGCTGTGGATGTGGTGTATCGCGGTGCGCGGTTCGAG 447

Db 371 CTGATCATGACCTGTTGAGCAACGACCGAGCTGGCGATCGGATCGGTGCGTCCAGC 430
QY 448 GATTCCAGCTGTTTCCCGCGCTCTGGGGCGGATGCCAATGGTGGTGGCGCCGCGAAA 507
Db 431 GATCCACCTGACCGCGCCGCTGGGGCGGAGCGGCTGCACATCTCTGCCAGCCGG 490
QY 508 AGCTATCTCACAAATACGAGCATACCGGAAACCCGCGGATTGAGTAGTCAATTCATGG 567
Db 491 GAGTACCTGCGCGCACGCGCGCGGAGCATCGAGGGCTGCGAGGGCACACCCCTG 550
QY 568 CTTGAATACAGGTGC 583
Db 551 CTGGCTTCACCGAGC 566

RESULT 15

US-09-489-039A-454
; Sequence 454, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 454
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-454

Query Match 11.1%; Score 103.2; DB 4; Length 933;
Best Local Similarity 49.3%; Pred. No. 4.8e-23;
Matches 270; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

QY 1 ATGGAAACGACTAAACGCAATGCGGTGTTTGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 28 ATGGACCGGTTTTCAGCTCTCAAGCCTTCACCCGCTTGTGAAAGCGCGAGCTTTACC 87
QY 61 GCGCGCGCCAGACAGATGAGCGTTTCGTCCATCAGTCAGACCGGTATCAAAACTG 120
Db 88 CGCGCGCGGAGTCGTGTAATATGCCAACGCCACCCCTCAGCAAAACCATTCAGCAACTG 147
QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTGACCGAA 180
Db 148 GAAGCGCATCTTGGGTTTCCCTCCTGACGCGCACCGCGCGTATTACCGTGACGCGG 207
QY 181 GCGGTAGAAATTTACTACAGGGCTGCGGTGATGCTTTCATGAAGTCAGGATGTTTCA 240
Db 208 GAAGGACGGGAGTACTATGAAAAGCCCGCTGCTGCTGGAAGACCTGGAAGAGATCGAC 267
QY 241 GAGCAACTGTATGCTTCAATAACACCCCGATCGGAGCGCTACGATTTGGCTGTTTCA 300
Db 268 GGTCTTTCAATACCGCCGCAATAGCGAGGCGCATCTACGATCGCATCGGCGGG 327
QY 301 ACTATGGCAAAATGTTCTCCCGGGGTGACAGCAAAATGCTGAAAGATPACCCAGGT 360
Db 328 TCGACCGGTGGGATGCTGATCGCGTGTGCGGGACTTTATGACCTCTGGCGGGAT 387
QY 361 TTGAGCGTCAATCTGGTTACCGGAATTCAGCGCCCGACCTGATTGCGGCGGTGGAT 420
Db 389 ATCCGTATGATTGAGGTGGCGGATAGCCCGGGACCTCATCAGCGGCAATATTGAC 447
QY 421 GTGGTATCCGGCTGGCGGCTTGAGGATTCAGGCTGTTTCCCGCGGCTCTGGGCGG 480
Db 448 TCGGCATCCGCGGGTCCGATGAGAGCTGACGCTGATCGCCCGTAAATCGCGAG 507
QY 481 ATGCCAATGGTGTGCGCCCGGAAAGAGCTATCTCACAAATACGGCATACCGGAAAAA 540

Db 508 GCGACGCTGGTCACTTGGCGCACCCCGGGCTATCTTCAGCGCTACGGCACCCCGCCTCG 567
QY 541 CCCGCCGA 548
Db 568 CCGGATGA 575

Search completed: December 18, 2004, 20:48:55
Job time : 108.036 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:42:54 ; Search time 645.306 Seconds
(without alignments)
7956.501 Million cell updates/sec

Title: US-10-759-889-1
Perfect score: 930
Sequence: 1 atggaacgactaaacgcacat.....ggcgcgggaagaagaagtaa 930

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4093002 seqs, 276041825 residues
Total number of hits satisfying chosen parameters: 8186004

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	930	100.0	930	17	US-10-759-889-1
2	112.4	12.1	939	16	US-10-282-122A-11996
3	109.6	11.8	945	16	US-10-282-122A-32141
4	105.6	11.4	936	16	US-10-282-122A-12996
5	99.8	10.7	918	15	US-10-127-032-41
6	96.4	10.4	939	16	US-10-282-122A-13932
7	93.6	10.1	945	16	US-10-282-122A-33170
8	93.6	10.1	954	16	US-10-282-122A-30384
9	89.8	9.7	896	16	US-10-282-122A-13983
10	78.4	8.4	936	16	US-10-282-122A-13759
11	72.6	7.8	1830121	14	US-10-329-960-1
12	72.6	7.8	1830121	16	US-10-329-670-1

C 13	72.6	7.8	1830121	18	US-10-158-865-1	Sequence 1, Appli
C 14	71.8	7.7	912	16	US-10-282-122A-13890	Sequence 13890, A
C 15	71	7.6	690	16	US-10-282-122A-11194	Sequence 11194, A
C 16	67	7.2	945	16	US-10-282-122A-7493	Sequence 7493, Ap
C 17	59.2	6.4	945	16	US-10-282-122A-11438	Sequence 11438, A
C 18	59	6.3	396	11	US-09-864-408A-2317	Sequence 2317, Ap
C 19	56.4	6.1	2731748	17	US-10-297-465A-1	Sequence 1, Appli
C 20	54.8	5.9	1074	18	US-10-425-115-98485	Sequence 98485, A
C 21	51.6	5.5	2409	15	US-10-085-959-56	Sequence 56, Appl
C 22	48.6	5.2	2731748	17	US-10-297-465A-1	Sequence 1, Appli
C 23	48.4	5.2	2547	16	US-10-264-237-933	Sequence 933, App
C 24	48	5.2	915	16	US-10-282-122A-30134	Sequence 30134, A
C 25	48	5.2	6133	14	US-10-114-170-15	Sequence 15, Appl
C 26	48	5.2	45613	16	US-10-672-787-22	Sequence 22, Appl
C 27	46.2	5.0	878	16	US-10-282-122A-23342	Sequence 23342, A
C 28	46.2	5.0	882	16	US-10-282-122A-36777	Sequence 36777, A
C 29	46	4.9	942	16	US-10-282-122A-39847	Sequence 39847, A
C 30	45	4.8	966	15	US-10-238-075-579	Sequence 579, App
C 31	45	4.8	2489	15	US-10-085-959-44	Sequence 44, Appl
C 32	45	4.8	2498	15	US-10-238-075-578	Sequence 578, App
C 33	44.4	4.8	942	16	US-10-282-122A-7496	Sequence 7496, App
C 34	44.2	4.8	903	15	US-10-238-075-619	Sequence 619, App
C 35	44.2	4.8	16057	15	US-10-085-959-189	Sequence 189, App
C 36	44.2	4.8	16066	15	US-10-238-075-610	Sequence 610, App
C 37	42.4	4.6	909	16	US-10-282-122A-31417	Sequence 31417, A
C 38	42.2	4.5	906	16	US-10-282-122A-13738	Sequence 13738, A
C 39	42	4.5	94750	16	US-10-672-787-38	Sequence 38, Appl
C 40	41.8	4.5	350	11	US-09-864-408A-8297	Sequence 8297, Ap
C 41	41.8	4.5	602	9	US-09-974-300-5894	Sequence 5894, Ap
C 42	41.8	4.5	2256646	17	US-10-470-565-1	Sequence 1, Appli
C 43	41.6	4.5	891	16	US-10-282-122A-31440	Sequence 31440, A
C 44	41.6	4.5	909	16	US-10-282-122A-40071	Sequence 40071, A
C 45	41.4	4.5	487	9	US-09-974-300-5831	Sequence 5831, Ap

ALIGNMENTS

RESULT 1
US-10-759-889-1
; Sequence 1, Application US/10759889
; Publication No. US20040157331A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Co., Inc.
; APPLICANT: Van Dyk, Tina K
; TITLE OF INVENTION: Regulator/Promoter for Tunable Gene Expression and Metabolite
; TITLE OF INVENTION: Sensing
; FILE REFERENCE: CL2036 US NA
; CURRENT APPLICATION NUMBER: US/10759,889
; PRIOR FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: 60/440,965
; PRIOR FILING DATE: 2003-01-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(930)
US-10-759-889-1

Query Match 100.0%; Score 930; DB 17; Length 930;
Best Local Similarity 100.0%; Pred. No. 1.6e-310;
Matches 930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGAACGACTAAACGCATGTCGGTGTTCGCCAAAGTACGTGCTGCTTTTACC	60
Db	1	ATGGAACGACTAAACGCATGTCGGTGTTCGCCAAAGTACGTGCTGCTTTTACC	60
Qy	61	GCCGCCGCCAGACGACTACAGATGAGCGTTTCGTCCATCAGTCAGCGTATCAAACTG	120

Db 61 GCGCGCCGACAGCTACAGATGAGCGCTTTCCTCCATCAGTCAGACGGTATCAAACTG 120
Qy 121 GAAGATGAGTTCAGGTAAGCTGTAAACCGTAGCACACGAGCAATGGCTGACCGAA 180
Db 121 GAAGATGAGTTCAGGTAAGCTGTAAACCGTAGCACACGAGCAATGGCTGACCGAA 180
Qy 181 GCGGTAGAAATTAACACAGGCTGCGCTGATCTTCATGAAGTCAGAGTTCAT 240
Db 181 GCGGTAGAAATTAACACAGGCTGCGCTGATCTTCATGAAGTCAGAGTTCAT 240
Qy 241 GAGCACTGTATGCTTCAATTAACACCGCCATCGGAGCGCTACGATGGCTGTCTTCA 300
Db 241 GAGCACTGTATGCTTCAATTAACACCGCCATCGGAGCGCTACGATGGCTGTCTTCA 300
Qy 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATCTGAAGAATACCAGGT 360
Db 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATCTGAAGAATACCAGGT 360
Qy 361 TTGAGCGTCAATCTGCTTACCGGAATTCAGCCCGGCTGATTCGCGAGCTGGAT 420
Db 361 TTGAGCGTCAATCTGCTTACCGGAATTCAGCCCGGCTGATTCGCGAGCTGGAT 420
Qy 421 GTGGTATCGCGTTCGCGGCTGTCAGGATTCAGGCTGCTTTTCCGCGCTGCGCG 480
Db 421 GTGGTATCGCGTTCGCGGCTGTCAGGATTCAGGCTGCTTTTCCGCGCTGCGCG 480
Qy 481 ATGCCAATGGTGTGCGCGGGAAGCTATCTCACAAATACGGATACCGGAAAT 540
Db 481 ATGCCAATGGTGTGCGCGGGAAGCTATCTCACAAATACGGATACCGGAAAT 540
Qy 541 CCGCGCGATTTGAGTAGTCAATTCATGCTTGAATACAGGCTGCGCGGCAATGAT 600
Db 541 CCGCGCGATTTGAGTAGTCAATTCATGCTTGAATACAGGCTGCGCGGCAATGAT 600
Qy 601 GAATGATCGACCGGAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
Db 601 GAATGATCGACCGGAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
Qy 661 AATGATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 AATGATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 CAGTCAGATCCCGCGGCTTATGCTTATATACCGGAAAGATTAAGCTGCGCTGAAG 840
Db 781 CAGTCAGATCCCGCGGCTTATGCTTATATACCGGAAAGATTAAGCTGCGCTGAAG 840
Qy 841 GTACAGTCTGATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 GTACAGTCTGATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 901 GAGATGACGCGCGCGGGAAGAGATTA 930
Db 901 GAGATGACGCGCGCGGGAAGAGATTA 930

RESULT 2
US-10-282-122A-11996
; Sequence 11996, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Orlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11996
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-282-122A-11996

Query Match 12.1%; Score 112.4; DB 16; Length 939;
Best Local Similarity 47.8%; Pred. No. 1.5e-27;
Matches 394; Conservative 0; Mismatches 421; Indels 9; Gaps 2;

Qy 1 ATGGAACGACTAAACGATGTCGGTGTTCGCAAGTAGTTCGAATTTGGCTCTTTTACC 60
Db 1 ATGGAACGACTAAACGATGTCGGTGTTCGCAAGTAGTTCGAATTTGGCTCTTTTACC 60
Qy 61 GCGCGCCGACAGCTACAGATGAGCGTTCCTCCATCAGTCAGAGGATCAAACTG 120
Db 61 GCGCGCCGACAGCTACAGATGAGCGTTCCTCCATCAGTCAGAGGATCAAACTG 120
Qy 121 GAAGATGAGTTCAGGTAAGCTGTAAACCGTAGCACACGAGCAATGGCTGACCGAA 180
Db 121 GAAGATGAGTTCAGGTAAGCTGTAAACCGTAGCACACGAGCAATGGCTGACCGAA 180
Qy 181 GCGGTAGAAATTAACACAGGCTGCGCTGATCTTCATGAAGTCAGAGTTCAT 240
Db 181 GCGGTAGAAATTAACACAGGCTGCGCTGATCTTCATGAAGTCAGAGTTCAT 240
Qy 241 GAGCACTGTATGCTTCAATTAACACCGCCATCGGAGCGCTACGATGGCTGTCTTCA 300
Db 241 GAGCACTGTATGCTTCAATTAACACCGCCATCGGAGCGCTACGATGGCTGTCTTCA 300
Qy 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATCTGAAGAATACCAGGT 360
Db 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATCTGAAGAATACCAGGT 360
Qy 361 TTGAGCGTCAATCTGCTTACCGGAATTCAGCCCGGCTGATTCGCGAGCTGGAT 420
Db 361 TTGAGCGTCAATCTGCTTACCGGAATTCAGCCCGGCTGATTCGCGAGCTGGAT 420
Qy 421 GTGGTATCGCGTTCGCGGCTGTCAGGATTCAGGCTGCTTTTCCGCGCTGCGCG 477
Db 421 GTGGTATCGCGTTCGCGGCTGTCAGGATTCAGGCTGCTTTTCCGCGCTGCGCG 480
Qy 478 GCGATGCCAATGTTGGTGTGCGCGGCAAGAGCTATCTCACAAATACGGATACCGGAA 537

```
Db 481 GAGAGCTACAGCGTCTGCGGATCGCGGGCTAGCTCGAGTCGCACGGCGTGGCGGAG 540
Qy 538 AACCCTCCGATTTGAGTGTATTCATCGCTTGAATACAGCGTGGG-----CCCGAC 591
Db 541 CGCGCGCGCGATCGCAGCAGCAGCTGTGCTGGGATGCTCGCGCGGGCTTTCACATTC 600
Qy 592 AATGAATTTGAAGTGTATCGCAGCGGAGGATCTCGACTCGCTGATCCCAAGGAAGA 651
Db 601 GACGAATGGCGCTGGCGGGCGGAAACGGCGAGAGTGTGCTCCGATCACGGCGCGCGG 660
Qy 652 TTGTGCTAATGATCCGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
Db 661 TTTCGCGTGAACGTGCGCGAGGCGCTGCGGTGGCGCTGCGGAGGGGATGGCGCTGGC 720
Qy 712 TACGTCGCGCTGATGCTGCTGATACGAGATCAATCTGCTGGGAGCTGAGATCTGCTG 771
Db 721 GGGCTCGCGCTGATTCGCGGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy 772 CGCGCTTACAGTACAGTCCAGCGCGCGGTTTATGCTTATATAC 815
Db 781 CCCGAGTACCGTGGCAGCTGATGACATCTACGGGCTGTATCC 824
```

RESULT 3

```
US-10-282-122A-32141
; Sequence 32141, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,309
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32141
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Pseudomonas putida
```

US-10-282-122A-32141

```
Query Match 11.8%; Score 109.6; DB 16; Length 945;
Best Local Similarity 50.7%; Pred. No. 1.4e-26;
Matches 290; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

Qy 1 ATGACGACGCTAAACCGCATGTCGTCGTTTGGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 1 ATGACACCCCTGCAAAACATGCGTGCTTTCAGTTGCGTAGCCCAACTCGCGAGCTTCACT 60
Qy 61 GCCCGCCCGACAGCTACAGATGAGCGTTTTCGTCATCATCAGTCAGAGCGTATCAAAACTG 120
Db 61 GCCGCTGCGGCGCACTGGATACGACACCGCAACGTCGCGGGCGGTCTCAACCTG 120
Qy 121 GAAATAGTGTGAGGTAAAGCTGTTAAACCGTAGACACGACGACATTCGCCCTACCGAA 180
Db 121 GAAGCCCATCTGCAAAACAGGCTGCTCAACCGTACCAACCGCGCATTTGCGCTGACCGAA 180
Qy 181 GCCCGTGAAGATTTACTACAGGCGTCCGTCGTATGCTTTCATGAAGTGCAGGATGTTTCA 240
Db 181 GCCGCGAAGCTTACCTGATGCTTTCGGAACAGATTTTACCTACGTCGGAAGAGCGGAG 240
Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGAGCGTACGCAATTGGCTGTCTTCTCA 300
Db 241 GCCGAGCCGACGACGCCCATGCGCGCGCGCGGCGAGTTGAAGTGCATTCGATGACT 300
Qy 301 ACTATGSCACAAATGTTCTCGCGCGGCTGACAGCCCAAAATGCTGAAGATATCCAGGT 360
Db 301 GGGTTCGCGCAGCATTCGTCGTGTCGATGCCATCGCCCGCTACCGCGAAACGCAACCGGAC 360
Qy 361 TTGAGCGTCAATCTGTTACCGGAATTCAGACCCCGACCTGATTCGCGACGGTCTGGAT 420
Db 361 GTGACCTTCGATTTGACCATGGCCACCGCGTCCCGATCTGCTCGACGAGGGTTATGAC 420
Qy 421 GTGTCGATC---CGCTCGCGCGGTTCAGGATTCAGAGTTCAGCGCTGTTTCCCGCGCTGGGC 477
Db 421 GTCTCATCTGTCGCGCACCGAACTGCGGACTCGGGGTTCGTTTCGACGCGCTGGGC 480
Qy 478 GCGATGCCAATGTCGTGTCGCGCGGAAAGATCTCTACAAATACGCAATACCGCATACCGAA 537
Db 481 ATCAGCTACAGCATTCGTCGTGCTGCGCGCTACATCGCCAGCATGCGCTGGCGGAC 540
Qy 538 AACCCTCCGATTTGAGTGTATTCATGCTTTCATGCT 569
Db 541 AAGCTCCCGACCTGCTCAAGCAGCGCTGCT 572

RESULT 4
US-10-282-122A-12996
; Sequence 12996, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIORITY FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

661 TTCCAGGTGAAGCTGGCCGGAAGCATGGCGTGGCGTATCGAGCGGCATGGCGCTGGG 720
712 TAGCTGCCCTGATGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTG 771
721 CTGATTCGATCTACTCGCGGATCAGCGCTTGGCAGCGGAGAACTGGTGTGGCTG 780
772 CCGCTTACAGTCAAGTCCAGCCCGGTTTATGCGTTATATACCGAAAAAGATAGCTG 831
781 CCGGAATACAGTCCGAGGAGATGAATCTGTACGGCTTGTACCCCTCGCGGAGTATCTG 840
832 CCGCTGAAGTACAGTCTGATCACTCGCTGACCGA 869
841 GACGGAAGATTGCGACGTGGATCGAGTTCTTCGCGCA 878

RESULT 5
US-10-127-032-41
; Sequence 41, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; PRIORITY FILING DATE: 2002-04-19
; PRIORITY FILING DATE: 2001-04-20
; PRIORITY FILING DATE: 2001-04-20
; PRIORITY FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-41

Query Match 10.7%; Score 99.8; DB 15; Length 918;
Best Local Similarity 49.4%; Pred. No. 3.3e-23;
Matches 288; Conservative 0; Mismatches 292; Indels 3; Gaps 1;

QY 2 TGNACGACTAAACCGCATGCTGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACCG 61
DB 14 TGGATCACCTCAACGCTCTCAAGGTCTTCGCGCGCGTGGCGCGAATGGCGGCTTCGCG 73
QY 62 CCGCGCGCAGACAGCTACAGATGAGCGTTTTCGTCTCATCATGATCAGCGGTATCAAACTGG 121
DB 74 CCGCGCGCGCGAGATGATCTCTCGCGCGCGCGCTGAGCAAGAACGCTCGCGAGCTGG 133
QY 122 AAGATGAGTTGCGAGTTAAGCTGTTAAACCGTAGCAGACCGAGCATGGCTGACCGAAG 181
DB 134 AAGCGCACCTCAAGGTGCGCTCTGATCAATCGCACCCCGCGAGCTGAGCTTACCGAGG 193
QY 182 CCGGTAGAAATTTACTACAGCGGTGCGCTGATGTTTCTATGAAGTGCAGGATGTTTCA 241
DB 194 CCGCGGAGTCTACCGCGAGCGCTGGAGGCGATCTCTGACGACCTCGAGCGCGCGAG 253
QY 242 AGCAACTGTATGCGCTTCAATAACACCCCGCATTCGCGAGCTTACGATGGCTGTTCTTCA 301
DB 254 CCGCGCTCACTTCGATCAGCAGCGCGCGCGCTGCTGCGGGTCAAGCGCGCGCTGTA 313
QY 302 CTATGGCACAATGTTCTCGCGCGGTGACAGCAAAATGCTGAAAGATATACCAAGTT 361
DB 314 CCCTCGCGCTTACCTGCTGACCCCGCGCATTCGCGCTTCTTCCAGCGTTATCCGAGC 373
QY 362 TGACGCTCAATCTGGTTTACCGGAATTCAGCCCGCGAGCTTGTATGCGCGCTCTGATG 421
DB 374 TGCCTGGTAACTGCTCTCTGAGGACCGCGCGCGAGCACTGATCGCGGAGGATCGACC 433

PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12996
LENGTH: 936
TYPE: DNA
ORGANISM: Burkholderia fungorum
US-10-282-122A-12996

Query Match 11.4%; Score 105.6; DB 16; Length 936;
Best Local Similarity 46.7%; Pred. No. 3.3e-25;
Matches 410; Conservative 0; Mismatches 459; Indels 9; Gaps 2;

QY 1 ATGACGAGCTAAACGCGATGCTGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 1 ATGATACGCTTCAAAACATGCGGTATTTTCGTTGCTGCGAGCGGAGCTTTACG 60
QY 61 GCGCGCGCCAGCAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGCTATCAAACTG 120
DB 61 GGTGCGCGCGAGCATCTGAACACGACACGCGCTATGCTGCGCGCGGTGCTCCGATCTG 120
QY 121 GAAGATGAGTTGCGAGTAAAGCTGTTTAAACCGTAGCAGCAGCAGCTGGCTGACCGAA 180
DB 121 GAAGCGCATCTGCGGACCGCGACTGCTGAACCGACCGACCGCGCGGATTCGCTGACCGAG 180
QY 181 GCGGTGAGAAATTTACTACAGCGCTGCGCTGATGCTTTCATGAAGTGCAGGATTTTCA 240
DB 181 GCGCGCGCGGCTATCTGAGCGCTGCGAGCAGATTTCTGCGTACGTGAGCAGCGGAA 240
QY 241 GAGCAACTGTATGCTTCAATACACCCCGATCGGACCGCTACGATGGCTGTTCTTCA 300
DB 241 GCGGAGCGCGGCGACCGCCATGCGCTGCGGCAAGCTCAAGGTCCAGCGCATGACG 300
QY 301 ACTATGCGCAAAATGTTCTCGCGCGGCTGACAGCCAAATGCTGAAAGATATCCAGGT 360
DB 301 AGTTTGGCGCAGCATATGTTGTCGCGCTGCTGCGCGCTATCAGCAGCGCTACCGGAC 360
QY 361 TTGAGCGCTCAATGCTGTTACCGGAAATTCAGCGCCCGCATGATTCGCGAGCGCTTGGAT 420
DB 361 GTCCATATCGAATGACGCTGCGGACGCGATGCGCGATCTGCTGACGAGGCTTCGAC 420
QY 421 GTGGTATC---CGCTGCGCGGTTGCGAGGATTCAGCGCTGTTTTCGCGCGCTGCGG 477
DB 421 GTGTCGTACCGTTCGCAACCGGTTTTCGCGGATTCGCGGCTGTTGTCGAGCGCTTCG 480
QY 478 GCGATGCAATGTTGTTGTCGCGCGGAAAGATATCTCACAATAACGCGATATACCGAA 537
DB 481 AGCGGCTTACGATTCGATGCGGCTGCGCGGCTATCTGAGCTGTCAGCGCGCTGCGAA 540
QY 538 AAACCGCGGATTTGAGTAGTCAATTCATGGCTGATATACAGGTGCGG-----CCCGAC 591
DB 541 ACGCCCGCGGACCTTTCGCGGCTACACTGCTGCAAAATGTTACGCGGTTTTCGCGAC 600
QY 592 AATGAATTTGAATTCGACCGGAGGAGGATCTCGACTGCGCTGATCCCAAGCAAGA 651
DB 601 GACAAATGAGACCTTCGAGCGCCCGACGCGGAGGAAACGATCGCCCTCGGACCGACT 660
QY 652 TTTGTGACTAATGATCGATGAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711

QY 422 TGGTATCCGCG---TCGGCGGCTTGCAGGATTCAGACCTGTTTCCCGCCGCTTGGCG 478
DB 434 TGGCCCTGCTGTAGCAGCCGCTTGCAGCTCCGGCTGTTCGGCCGCTGCTGG 493
QY 479 CGATGCCAATGCTGTGTCGGCGGGAAGCTATCTCACAATACGGCATACCGGAAA 538
DB 494 TCTTGAGACGCTGCTCTCGGGGCGCGCTTAOCTCAGTCAGATGGCCAGCGCTGC 553
QY 539 AACCCCGCGATTGTAGTAGTCAATTCATGGCTTGAATACAGCGT 581
DB 554 GCGCGAGGCCCTGCGCGACAGGATGATCCGCTTACGCT 596

RESULT 6

US-10-282-122A-13932
; Sequence 13932, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELIUPA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13932
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-13932

Query Match 10.4%; Score 96.4; DB 16; Length 939;
Best Local Similarity 45.5%; Pred No. 5, 1e-22;
Matches 426; Conservative 0; Mismatches 50; Indels 9; Gaps 2;

QY 1 ATGGAACGACTAAACCGCATGCTGGTGTTCGCAAGTAGTGAATTTGGCTCTTTTACC 60
DB 1 ATGGAACGCTCCAAACCATGCGCTGTTCGCGCGTGGTCCGATGCGGCGAGTTTACG 60
QY 61 GCGCGCGGACGACGACTACAGTACGCTTTTCGTCATCAGTCAGACGATCAAAACTG 120

DB 61 GCGCGCGGACGAGCTCAATTGACGACCGGCTACGATCGCGCGGCTGTCGATCTC 120
QY 121 GAAGATGAGTTGAGGTAAGAGCTGTTAAACGTAAGACACGACGATATGGCGCTGACCGAA 180
DB 121 GAGCGGCACTTGGCGACACGCGCTTCTGACCGGACGACGCGGCGCATTCGCGCTCACGAG 180
QY 181 GCGCGTAGAATTTACTACGAGGCTGCGCTGATGCTTTCATGAAGTACAGGATGTTTCA 240
DB 181 GCGCGGAGCGTTTATCTGACGCTGCGAGCAGATCTTCGCTACGTCGAGCAGCGGAG 240
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGATTTGGCTGCTTTTCA 300
DB 241 GCGGAGGCAAGCGACGCGCGCGCGCTCCGCAAGCTGAAAGTGCACTGCATGACG 300
QY 301 ACTATGGCACAAAATGTTCTCGCGGCTGACAGCCAAAATCTGAAAGATAACCCAGGT 360
DB 301 AGCTTCGGCAGCACTACACGCTGCGCGGCTGCGGCTACCGGACGCGCTATCCGAA 360
QY 361 TTGAGCGTCAATCTGGTTACCGGAAATTCAGCCCCCGACCTGATTTGCCGACGCTTGGAT 420
DB 361 GTCCAGGTCGACCTGACGCTCGCGAGCGGCTCGGATCTGCTGACGAGGCTTCGAC 420
QY 421 GTGCTGATCCGCTCGCG---GCGTTGACAGATTCAGCCTGTTTCCCGCGCTTGGGC 477
DB 421 GTGCTGCTGCTGCTCGCGCGGAGCTGCGCGATTCGCGGCTCGTGTGCGACGCGCTCGC 480
QY 478 GCGATGCCAATGCTGCTGCGCGGGAAGGCTATCTCACAATACGACATACCGGAA 537
DB 481 GAGAGCTTCAGCTGCTGCTCGCGGCTCGCGGCTACATCGAGCAGCAGCGCGCGCAG 540
QY 538 AACCCTCCGATTTGAGTAGTATTAATGCTTGAATACAGCTG-----CGCGCGAC 591
DB 541 CGCGCGCAGGATCTTGGCGGCGGCTGCTGCTCGGATGCTGCTCGCGCGGCTGTACTGG 600
QY 592 AATGAATTTCAATGATCGCGAGGAGGATCTCGACTCGCGCTGATCCACAGGAAGA 651
DB 501 GACGAATGGAAGCTGTTGGGCGCGAGGCGACAGCGTACGCTCGCGCGCGCGCG 660
QY 652 TTTGTGACTAATGATCGATGACGCTGCTGCTGCTGCTGACGCGGCTGCGGAGTCCG 711
DB 661 TTTGCGGTGAACGCTGCGCGAGGCGCTGCGCGCGGCTGCGAGGCGCATCGCG 720
QY 712 TACGTGCGCTGATGTTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTG 771
DB 721 GTGCTGCGCTCTATTGCGGATCGCGGCTGCGCAAGCGCGATTTGTCGCGCGTATG 780
QY 772 CCGGTTTACGAGTACAGATCCACCGCGGTTTATGCTTATATACCGAAAAGATAAGCTG 831
DB 781 CCGGAATACCGGTCGCGACGCTGATGAACATCTACGCGCTCTACGCTCGCGCGATCTC 840
QY 832 CCGCTGAAGGTACAGTCTGATCACTCGCTGACGAGATTTTGTGAGGTTCGTA 891
DB 841 GACGCGAAGATCCGACGCTGGGTGCTGATTTTCTGCGGACGAGTTGCGCTGATCTCGAA 900
QY 892 TTGTTTTCAGGAGATGACGCGCGCGGGAAGAGAG 927
DB 901 GCGGACGAGGCGGCTGCGAGCGCTTCACCGTGCAG 936

RESULT 7

US-10-282-122A-33170
; Sequence 33170, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert		
; APPLICANT: Forsyth, R.		
; APPLICANT: Xu, H.		
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms		
; FILE REFERENCE: ELITRA.034A		
; CURRENT FILING DATE: 2003-02-20		
; PRIOR APPLICATION NUMBER: 60/191,078		
; PRIOR FILING DATE: 2000-03-21		
; PRIOR APPLICATION NUMBER: 60/206,848		
; PRIOR FILING DATE: 2000-05-23		
; PRIOR APPLICATION NUMBER: 60/207,727		
; PRIOR FILING DATE: 2000-05-26		
; PRIOR APPLICATION NUMBER: 60/230,335		
; PRIOR FILING DATE: 2000-11-27		
; PRIOR APPLICATION NUMBER: 60/257,931		
; PRIOR FILING DATE: 2000-12-22		
; PRIOR APPLICATION NUMBER: 60/267,636		
; PRIOR FILING DATE: 2001-02-09		
; PRIOR APPLICATION NUMBER: 60/269,308		
; PRIOR FILING DATE: 2001-02-16		
; Remaining Prior Application data removed - See File Wrapper or PALM.		
; NUMBER OF SEQ ID NOS: 78614		
; SOFTWARE: Patentin version 3.1		
; SEQ ID NO 33170		
; LENGTH: 945		
; TYPE: DNA		
; ORGANISM: Pseudomonas syringae		
; US-10-282-122A-33170		
Query Match		10.1%; Score 93.6; DB 16; Length 945;
Best Local Similarity		49.0%; Pred. No. 4.7e-21;
Matches	280; Conservative	0; Mismatches 289; Indels 3; Gaps 1;
QY	1	ATGGAAGCACTAAACCGATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 60
DB	1	ATGGACACCTTGCAAAACATGCGGCTTTACGACGCTGGACAGCGCGGAGCTTCACC 60
QY	61	GCGCGCGCGACAGACTACAGATGAGCGTTTCGTCATCAGTCAGCGGTATCAAAACTG 120
DB	61	GCGCGCGCTGCGGTGCTGGACACCAACGACCGCAACGCTCGCGCGCTGTGTCCAACTT 120
QY	121	GAAGTAGTTGCGAGTAAAGCTGTTAAACCGTAGCAGCGATTTGGCTGGAGAGCCGAA 180
DB	121	GAAGCGACCTGCAAGCCCGGCTGCTAACCGCACCAACGACGCAATGTTCTTACCAGAA 180
QY	181	GCGGTGAGAAATTTACTACAGGCGTGCCTGTATGCTTTCATGAAGTCGAGGATGTTTCA 240
DB	181	GCGGCAAGCGTTTACTCTGTTGCGTGCGAGCAGATTTCTGGCTCGGTGAGGAAGCCGAG 240
QY	241	GAGCAACTGTATGCTTCAATAAACCCCAATCGGACGCTACGCAATTTGGCTGCTTCTCA 300
DB	241	GCGGAAGCGAGCGAGCCCATGCAAGGCCCGCGGAGCTCAAGGTGCAATTTGATGCCG 300
QY	301	ACTATGGCACAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360
DB	301	GGTGTGCGTCAGCACTATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY	361	TTGACGCTCAATCTGTTACCGGAATTCACGCCCGACCTGATTTGCCAGCGTCTGGAT 420
DB	361	GTGCGCTTCGACCTGACCATGACCCACCGGTGCGAGACTTGTCTGGAAGAGGCTTCGAC 420
QY	421	GTGGTGATC---CGGCTCGGCGGTGAGGATTTCCAGCTGTTTTCGCGCGCTGGGC 477
DB	421	GTGTGATCGTGTGCGGCAGCGAACTGCGCGATTTGGGCTTTGTGTGCGAAGCGCTGGGC 480
QY	478	GCGATGCCAATGGTGTGCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAA 537

Db	481	ATTACCTACAGCATGCGTGTGTCATCCCCCGAGTAGTCAAAACCTTCGGCATGGCCAC 540
QY	538	AAACCCCGCGATTTCAGTAGTCAATTCATGCGCT 569
DB	541	AAGCCCGCAGACCTTTTGAACCATGCGTGCCT 572

RESULT 8

US-10-282-122A-30384

; Sequence 30384, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 30384

; LENGTH: 954

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-10-282-122A-30384

Query Match

10.1%; Score 93.6; DB 16; Length 954;

Best Local Similarity

49.0%; Pred. No. 4.8e-21;

Matches

280; Conservative

0; Mismatches 289; Indels 3; Gaps 1;

QY 1 | ATGGAAGCACTAAACCGATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 60 || DB | 1 | ATGGACACCTTGCAAAACATGCGGCTTTTCGTCATCAGTCACGCGTATCAAAACTG 120 |
QY	61	GCGCGCGCGACAGCTACAGATGAGCGTTTCGTCATCAGTCAGCGGTATCAAAACTG 120
DB	61	GCGGCGTCCAGCAGCTCAACACACCGCTATGTTTCGCGGCGGTGGCCAACTC 120
QY	121	GAAGTAGTTGCGAGTAAAGCTGTTAAACCGTAGCAGCGATTTGGCTGGAGAGCCGAA 180
DB	121	GAAGCGACCTGCAAGCCCGGCTGCTAACCGCACCAACGACGCAATGTTCTTACCAGAA 180
QY	181	GCGGTGAGAAATTTACTACAGGCGTGCCTGTATGCTTTCATGAAGTCGAGGATGTTTCA 240
DB	181	GCGGCAAGCGTTTACTCTGTTGCGTGCGAGCAGATTTCTGGCTCGGTGAGGAAGCCGAG 240
QY	241	GAGCAACTGTATGCTTCAATAAACCCCAATCGGACGCTACGCAATTTGGCTGCTTCTCA 300
DB	241	GCGGAAGCGAGCGAGCCCATGCAAGGCCCGCGGAGCTCAAGGTGCAATTTGATGCCG 300
QY	301	ACTATGGCACAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360
DB	301	GGTGTGCGTCAGCACTATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY	361	TTGACGCTCAATCTGTTACCGGAATTCACGCCCGACCTGATTTGCCAGCGTCTGGAT 420
DB	361	GTGCGCTTCGACCTGACCATGACCCACCGGTGCGAGACTTGTCTGGAAGAGGCTTCGAC 420
QY	421	GTGGTGATC---CGGCTCGGCGGTGAGGATTTCCAGCTGTTTTCGCGCGCTGGGC 477
DB	421	GTGTGATCGTGTGCGGCAGCGAACTGCGCGATTTGGGCTTTGTGTGCGAAGCGCTGGGC 480
QY	478	GCGATGCCAATGGTGTGCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAA 180

QY 181 GCGGCTAGATTTACTACAGGCTGCGCTGATGCTTTCATGAGTGCAGATGTTTCAT 240
Db 181 GCGGCTAGGCTTACTGCTGCGCTGCGAGAGATCTTCGCTGCTGAGGAGCCGAG 240
QY 241 GAGCAACTGTATGCTTCAATTAACACCCCGGCTGAGAGAGCTGCTGAGTGTGTTTCA 300
Db 241 GCGGAAGCCAGGACGCGCATGCCGACCCGCGCGCAAGCTGAAGCTGCATTGCATGACC 300
QY 301 ACTATGGGACAAATGTTCTGCGCGGCTGAGAGCAAAATGCTGAAAGATATCCAGGT 360
Db 301 GGCATCGGCGAGCATACGTGATCGCGCATCGCGGCTACCGCCAGCATACCGGAC 360
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCCCGACCTGATGCGCGAGGCTCTGGAT 420
Db 361 GTGACCTCGACCTGACCATGCCAACCGCGTCCGCGACCTGCTCGAGAGGCTTCGAC 420
QY 421 GTGCTGATCCGCT---CGGCGGTTGAGAGATTCACGCTGTTTTCGCGCTGCGGC 477
Db 421 GTGGCCATCGTGGTGGCTCCGAGCTGCCGATTCGCGCTTCATTTCCAGCGCATCGGC 480
QY 478 GCGATGCCAATGTTGTTGCGCGCGGAAAGCTATCTCACACATACGGCATACCGGAA 537
Db 481 GAGACTACAGCATCTTTTGGCCCTCGCGGAGTACCTGCGCTGAAAGCGGACCGCGAGC 540
QY 538 AAACCGCGGATTTGAGTAGTCAITTCATGGCT 569
Db 541 GAGCCTTCGAGCTGGCCAGCCAGCTGCCT 572

RESULT 9
US-10-282-122A-19983
; Sequence 19983, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19983
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
; US-10-282-122A-19983

Query Match 9.7%; Score 89.8; DB 16; Length 896;
Best Local Similarity 47.2%; Pred. No. 9.5e-20;
Matches 342; Conservative 0; Mismatches 377; Indels 6; Gaps 2;

QY 49 GGCTTTTACCOCGCCGACAGCTACAGATGAGGTTTTCGTCATCAGTCAAGC 108
Db 40 GGCAGCTTTTCGCCCGCGCGGAGGCTGGCATGTCCGCCACGCGGTAGCAACGCC 99
QY 109 GTATCAAACTGAAGATGAGTTGCAGGTAAAGCTGTTAAACGCTAGCACACGACAT 168
Db 100 ATCGCGGGCTGGAGAGCCGCTTGTATCCCGCTTTTAAACGACACCCGCGGTG 159
QY 169 GGCTGACCAAGCCGCTAGAAATTTACTACAGGCTGCGCTGCTATGTCTCAATGAAGTG 228
Db 160 GCGCTCACCGACCGGACAGCGCTAGCTGGCGGTATCGGCCGCGCTGCAGGAGATC 219
QY 229 CAGGATGTTTATGAGCAACTGTATGCTTCAATAACACCCCATCGGAGCTACGCAIT 288
Db 220 CGTCTCGCCAGGAGATGCCACAGCGACACCGGGGAAACCCCGCGGACGCTCGCTG 279
QY 289 GGCTGTTCTTCAACTATGGCACAATAATGTTCTCGCGGGCTGACAGCAAAATGCTGAA 348
Db 280 AAGTGCAGCACTATCGCGGCTGTTTCTGACAGCTGCTGATGACTTATGATC 339
QY 349 GAATACCCAGTTTGAAGCTCAATCTGTTACCGAATTCACGCCCGGACCTGATGCC 408
Db 340 CGCTACCCGAAGATGCGCGTAGAGACCGTCAAGCAAGCGCAATGATCGCATCGTCGG 399
QY 409 GACGCTCTGATGTTGATCGCGCTGCGCGCTTCAGGATTCAGGCTGTTTTCGCGC 468
Db 400 GAAGCTACGACCGGGGATCCGCTTGAAGATTCGCTGCTCAGGACATGATCGCGTG 459
QY 469 CGTCTG---GGCGGATGCGCAATGTTGTTGTCGCCGCGGAAAGCTATCTCAACAATAC 525
Db 460 CCGCTGACCGGGGAGATCCGCCAGTTGTTGTCGCCGCGGACTATTTTGCACGCCAC 519
QY 526 GGCATACCGGAAAAACCCGCGGATTTGAGTAGTCAATTCATGCTTGAATACAGCGTGG 585
Db 520 GGCATTCGCGAGCGCGGAGATCTGTTTTCGATCAGGGGATGGCATGCGCATGGCC 579
QY 586 CCGGAC---AATCAATTTGAATCTGATCGCACCGGAGGATCTCGACTCGCCTGATCCCA 642
Db 580 CACGGCGGATCTACCTTTGGGAGCTGCGCGCGGGAACGTAACGCCCTTTCGCGTG 639
QY 643 CAAGGAAGATTTGATTAATGATCCGATGACGCTGCTGCTGCTGACGCGCGGTGCC 702
Db 640 CCGCGCGCTTTCGAGCTCCGATCTTTTTCGCTCGATCCGCGCGGTAAAGCGGGATTA 699
QY 703 GGGATCGCTACGTCGCGCTGATGGGTGATCAACGAGATCAATCTGTTGGGAGCTGGAG 762
Db 700 GGGGTGGATTTTTCGCGGACTGTATATCCGGGAAGAGCTGAGCGTGGAGAGCTGGTG 759
QY 763 ATCCT 767
Db 760 AGCGT 764

RESULT 10
US-10-282-122A-13759
; Sequence 13759, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13759
LENGTH: 936
TYPE: DNA
ORGANISM: Burkholderia fungorum
US-10-282-122A-13759

Query Match 8.4%; Score 78.4; DB 16; Length 936;
Best Local Similarity 47.2%; Pred. No. 8.6e-16;
Matches 272; Conservative 0; Mismatches 301; Indels 3; Gaps 1;
QY 1 ATGGACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 1 ATGGACACATTGGAGACATCGGATGTTTCGTTCGAGTCGTGGAAGCAGCGAGCTTCACG 60
QY 61 GCGCGCGCAGACAGCTACAGTAGCGGTTTGTTCATCAGTCAGACGGTATCAAACTG 120
DB 61 AAGGCGCGATCTCTTTTGTTCGTCGCGACACCTCAGGTATCGCGTCCCATACAGAACTG 120
QY 121 GAAGATGAGTTCGAGTAAAGCTGTAAACCGTACACAGCAGCATTTGGCTGACCGAA 180
DB 121 GAGTCGGATTTCGATCCCGTCTGTAAACCGGACGACGACATGTCTACTAACCGAA 180
QY 181 GCGGTAGAAATTTACTACAGGGCTGCGGTGTATGCTTTCATGAAGTCAGGATTTTCAT 240
DB 181 GCGGCGGAACGTTACTGCAAGTTTGGCAACAGATTTCTTGAATACGTGCGTCTGCGCGAA 240
QY 241 GAGCACTGTATGCTTCAATAACACCCCTCGGAGCGCTACGCATTTGGCTGTTCTTCA 300
DB 241 GCGGAGCGGAGGAGCGGCGCGAGCGGTCGCGGAAAGCTGCGGATTCACGACACACC 300
QY 301 ACTATGGCAAAATGTTCTCGCGGGGTGACAGCAAAATGCTGAAGAATATCCAGGT 360
DB 301 AGTTTCGGGCAACATTAATCTGCGGCACTCATCGCGGCTATAACACGAGATTCGCGAC 360
QY 361 TTGAGCGTCAATCTCGTTACCGGATTCAGCCCCGACCTGATTGCGGAGGTTCTGAT 420
DB 361 GTCTCTATGATCTTGTACTCTGCAACAGGAGTGGCGGACATGATCGAGGAGGTTTGTAT 420

QY 421 GT---GGTATCGCGTCGGCGGTTGACAGATTCCAGGCTGTTTCCCGCGCTGGGC 477
DB 421 GTATCGGTGTCGTGCTAGAGAAATTGGCGGATTCGGCACTCGTCTCACAACGCAATCGT 480
QY 478 GCGATGCCAATGTCGTGTCGCGCGGAAAAAGTATCTCACACAATACCGCATACCGGAA 537
DB 481 ACCACGTGCACCTGTTCTCTGTGCTTCGCCCCAGTACCTTCGAAGCGAGGTATGCTGTC 540
QY 538 AAACCCGCGGATTTGAGTAGTCTCATTTATGCTTGAA 573
DB 541 TCACTAGAAGATCTGAACACGATACGTCGCTGCAA 576
RESULT 11
US-10-329-960-1/c
Sequence 1, Application US/10329960
Publication No. US20030099277A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Pragn
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10/329,960
CURRENT FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match 7.8%; Score 72.6; DB 14; Length 1830121;
Best Local Similarity 45.7%; Pred. No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
Query Match 7.8%; Score 72.6; DB 16; Length 1830121;
Best Local Similarity 45.7%; Pred. No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;
Qy 1 ATGGAACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTCGAATTTGGCTCTTTTACC 60
Db 1450229 ATGGATAAAGCTCAACGCGATTTCTATTCTGCAAGTAGTTCGAACCCCAAGTTTACT 1450170
Qy 61 GCGCGCCGACAGAGTACAGTGGGTTTCCTCCATCATGAGCGGTATCAAACTG 120
Db 1450169 TTGGCAGCAAGCAACAAATATTTCTGCGCATGGCAAGCAATTAGTTTCAATTA 1450110
Qy 121 GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCACACGACGATTCGCTGACCGAA 180
Db 1450109 GAAGAACATTTAAACACACGATTGTTACACGTTACACACGAGAAAAATATGCTACCGAA 1450050
Qy 181 GCGGTAGATTTACTACGAGGCTGCGTCTGATGCTTCATGAGTGCAGGATGTTCTAT 240
Db 1450049 GCTGGAATGATGATTATATCAACGCTGTCAAGGCAATTTTGCTTGAATGAGCAGAT 1449990
Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGCTGTCTCA 300
Db 1449989 TCAGCAATTAATGCTTACCACTTCAATACAGGAAATTTACTGATTTCCGTTCTCGT 1449930
Qy 301 ACTATGGCACAATAATGTTTCGCGGGCTGACAGCCAAAATGCTGAAAGATACCCAGGT 360
Db 1449929 GATTTGGTTTACTTTTATTTGACCAAAATTTACCCACTTTTATGGCTAAGCATCTCAT 1449870
Qy 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGGACCTGATTCCGACGCTCTGGAT 420
Db 1449869 TTGCATATGAGTTGAATTTAATGATGAAATGATTTATATTCGAGAGCTATGAT 1449810
Qy 421 GTGATGATCGCGCTGCGCGGTTTCGAGGATTCAGGCTGTTTTCCGCGCTCTGGCGCG 480
Db 1449809 CTTGCTCTCGTATGCTTATATGGAAGATAGCTCGTTGGTTTACGTAATAATAGGAAT 1449750
Qy 481 ATGCCAATGTTGTCGCGCGGAAAGATATCTCACAAATACGGCATACCGGAAAAA 540
Db 1449749 ACTACTGTTCAATTTGAGCGCTGCGCTAATATTCTTGAACGAATGATATCCACAACG 1449690
Qy 541 CCGCGCGGATTT 551
Db 1449689 CTTGATGATTT 1449679
RESULT 13
US-10-158-865-1/c

```
Sequence 1, Application US/10158865
Publication No. US20040203093A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
FILE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB186P2C1D1
CURRENT APPLICATION NUMBER: US/10/158,865
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 09/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
```



```
/ NAME/KEY: misc feature
/ LOCATION: (122167)..(122167)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (122336)..(122336)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131340)..(131340)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131360)..(131360)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (139910)..(139910)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (140398)..(140398)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (142750)..(142750)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145058)..(145058)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145171)..(145171)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145942)..(145942)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147197)..(147197)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (150841)..(150841)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152500)..(152500)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152530)..(152530)
```

```
Query Match 7.8%; Score 72.6; DB 18; Length 1830121;
Best Local Similarity 45.7%; Pred No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGATGCGGTGTTGCCAAGTAGTGTGAATTTGGCTCTTTTACC 60
DB 1450229 ATGGATAAAGTCAACGCGATTTCTATTCTGCAAGTCAITGAAACCAAGTTTACT 1450170
QY 61 GCGCGCGGACAGCTACAGATGAGCGTTTGTCCATCAGTCAGCGGTATCAAACTG 120
DB 1450169 TTGGCAGCAACGCAAAATATTTCTGTGGCGATGGCAAGCAATTTAGTTTCAATTA 1450110
QY 121 GAAGATGAGTTGAGTAAAGCTGTTAAACCGTAGCACACGCAATTTGGCTGACCGAA 180
DB 1450109 GAAGAACATTTAAACACGATTTGTACAACTACCAAGAAATTTGCTTACCGAA 1450050
QY 181 GCCGTGAGATTTACTACAGCGGTGCGGTGATGCTTCTATGAGTCAGCGATGTTCAAT 240
```

```
DB 1450049 GCTGGAATGATGATTATTAACCGCTGTCAAGGCATTTTGTGTTAGTAGAAGCAGAT 1449990
QY 241 GAGCAACTGTATGCCITTCATTAACACCCCATCGGAGCGGTACGATTTGGCTGTCTTCA 300
DB 1449989 TCCAGCATTAATCTCAGCTTACCAGTTTCAATTAAGGAAATTTACTGATTTCCGTTCGT 1449930
QY 301 ACTATGCCACAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGT 360
DB 1449929 GATTTTGGTTACTTTTATGACCAAAATTTACCCACTTTTATGGCTAAGCATTCCTCAT 1449870
QY 361 TTGAGCGTCAATCTGTTTACCGGAAATTCAGCCCCGACCTGATTTGCCAGCGTCTTGAT 420
DB 1449869 TTGCATATTGAAGTTGAATTAATGATAAGAAAATCGATTTATTTCCGAAGGCTATGAT 1449810
QY 421 GTGGTGATCCGCTCGCGGGCTTGCAGGATTCAGCGCTGTTTCCCGCCGCTCTGGGCGG 480
DB 1449809 CTTGCTCTTCTGTTTATGGAAGATGAGTCTGTTGTTTCACTAAATAGGAAC 1449750
QY 481 ATGCCAATGTTGTTGCGCGGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAAA 540
DB 1449749 ACTACTGTTCAATTTGCGAGCGTGCCTAATTAATCTTGAAACGATGGTATTCACAAACG 1449690
QY 541 CCGCGCGATTT 551
DB 1449689 CCGTATGATTT 1449679
```

RESULT 14

```
US-10-282-122A-13890
; Sequence 13890, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13890
; LENGTH: 912
```

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITEA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11194
LENGTH: 690
TYPE: DNA
ORGANISM: Burkholderia cepacia
US-10-282-122A-11194

Query Match 7.6%; Score 71; DB 16; Length 690;
Best Local Similarity 46.1%; Pred. No. 2.7e-13;
Matches 276; Conservative 0; Mismatches 320; Indels 3; Gaps 1;

Qy 1 ATGAACGACTAAACGCGATGCGGTGTTGCCAAGTAGTTGAATTTGGCTCTTTTACGCC 60
Db 1 ATGGATCTGATTTCAAAAGCATGCGGCTTCGTACCGTCGCGACGCGGCGACCTTCACC 60

Qy 61 GCCCGCGCAGACAGCTACAGATGACGGTTTGTCTCATCAGTCAGACGGTATCAAACTG 120
Db 61 GAGACCGGCAACCGCTTCGGCATCAGCGCGCGCTGCGCGCTCGCTC 120

Qy 121 GAAGATGAGTTGAGGTAAAGCTTTAAACCGTAGCAGCAGCATTTGGCTGACCGAA 180
Db 121 GAGCATCTCTCGGCATCCGGTTGCTGAACCGCAGCAGCGTTTCGATGTGCTGACCGAA 180

Qy 181 GCCGGTAGAATTTACTACAGGGCTCGCTTCATGAGTGCAGGATGTCAT 240
Db 181 GCGGCGAGCGCTATCTGGCGCGCGCGCGGACATCACGCTCGCGCTCGATGCGAGGAG 240

Qy 241 GAGCAACTGTATGCTTCAATAACACCCCATTCGGACCGCTACGCAATTTGGCTTTTCA 300
Db 241 CGGAGCGCGCGACGCGCGCGCTGTCATCGCGCGCGGCTGCGCGCGCATTTGCGAGCG 300

Qy 301 ACTATGCGCAAAATGTTCTCGCGCGGCTGACAGCCAAATGCTGAAGAAATACCCAGGT 360
Db 301 AGCATCGCAAAACCATTTCTGATGATTCGCTGATTCGCGCGGCTTACCGCGGAC 360

TYPE: DNA
ORGANISM: Burkholderia mallei
US-10-282-122A-13890

Query Match 7.7%; Score 71.8; DB 16; Length 912;
Best Local Similarity 45.1%; Pred. No. 1.6e-13;
Matches 393; Conservative 0; Mismatches 467; Indels 11; Gaps 3;

Qy 4 GAACGACTAAACGCGATGCGGTGTTGCCAAGTAGTTGAATTTGGCTCTTTTACGCC 63
Db 10 GATCGCTCGCGCGCTGCGCATTTATGAGCGTCTGCGCAGCGCAATTTACCGGT 69

Qy 64 GCGCGCAGACACTACAGATGAGCTTTCGTCCATCAGTCAGCGGTATCAAACTGAA 123
Db 70 GCGCGCGCGCGCTCGCTTTGAGCGCGCGCGGTGAGCTTGGCATCGGCACTCGAA 129

Qy 124 GATGATTTGAGTTAAAGCTGTTAAACCGTAGCAGCAGCATTTGGCTGACCGAAGCC 183
Db 130 AGCGAGCTGAAAGTGAAGCTTTCAGCGCAGCAGCGCGCGGTGAGCTGACGGAAGCC 189

Qy 184 GGTAGATTTACTACAGCGGCTGCGTGTATGCTTCTGATGAGTCAGAGTTTCATGAG 243
Db 190 GCGCAGCGCTATTACGACGCAACGCGCAGCGCTACCGGCGAGTCTGCGAGCGCGCAC 249

Qy 244 CAACTGTATGCTTCAATAACACCCCGCTCGGAGCGCTACGATTTGGCTGTTCTTCACT 303
Db 250 GCGCTGAACGACGCGCGCGCAACCGGAGGCGCTTTCGGATCAGCGCGCTCACCTC 309

Qy 304 ATGGCACAATAATGTTCTCGCGCGGTGACAGCGCAAAATGCTGAAGAAATACCGAGTTG 363
Db 310 GCGAAATCGCTGTCAGCGCGCGCGTGTGCGCGAGTTCTTCCGCGCTACCGAAGTG 369

Qy 364 AGCTCAATCTGTTTACCGGAATTCAGCGCCCGACCTGATTCGCGACGGTCTGATGTG 423
Db 370 AGCTCGAGATCGCTACAGAGATCATCTGTCGATCGTCAAGGACGCTCGACGCC 429

Qy 424 GTGATCCGCTCGCGCGGTGAGGATTCAGCGCTGTTTCCGCGCTGCGGCGCGATG 483
Db 430 GGAATCCGCTCGCGACCGCGTGCAGCGCGGATGTTGGGCTGAAGATCGCGCGCGC 489

Qy 484 CCAATGCTGTGCGCGCGAA---AGCTATCTCACATATACGATACCGGAA 540
Db 490 CTCAGTGCCTGCTGCGCGCGCGCGCTATCTCGCGCAGCAGCGGTGCGCGCGAG 549

Qy 541 CCGCGGATTTGAGTAGTATTCATGCTTGAATACAGCGTGC-----GCGCGCAAA 593
Db 550 CTCGCGATCTCGCGCGCACATGTCATCGCTTTCGTAAGAGCGCGCGCTG 609

Qy 594 TGAATTTGAATGATCGCACCGAAGGATCTCGACTGCTGATCCCAAGGAGATTT 653
Db 610 CACAAATGCGCTGCGCGACGCGCGCGCGCGCTGAGCTGCGCTGCGCGCGCTT 668

Qy 654 TGTGACTAATGATCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 713
Db 669 GGTACACCGATACGCGCGCGCTGATCGACCGCGCGCGCGCGCGCGCATCGCTT 728

Qy 714 CFTGCGCTGATGTGGGTGATCAACAGATCAATCGTGGGAGCTGAGATCTCTGCTGCC 773
Db 729 CGTGTTCATGCGCGAGCGCATCGAGACGACCTGCGCGCGCGCGCGCGAGTGTCT 788

Qy 774 GGTATACAGTACAGTCCAGCGCGCTTTATGCTTATATACGAAAGAAATAGCTGCC 833
Db 789 GCGGCGCGGTGCGGAGCGCTGCGCGCGATGCTTACTAGCGGAACCGCAAGCAGT 848

Qy 834 GTTGAAGGTACAGGTCTGATCAACTCGCTG 864
Db 849 GCGCGCGAAGTGTGCGCGCTTATCGCGGTG 879

```
QY 361 TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCCCGACCTGATTCGCCGACGGTCTGGAT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GTGAGCGTCAATCTGGTTACCGGAATTCAGCCCCCGACCTGATTCGCCGACGGTCTGGAT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GT---GGTGATCCGCGTGGCGGTTGCGAGATTCCAGGCTGTTTTCCCGCGCTCTGGGC 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GTGCGGTGATCGCGATGCGGACCTGCCGAGTTCGAGGAGGTCCGATCGACGTCGGC 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 GCGATGCCAATGGTGTGGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAA 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 CGGATCCGACGCTGCTGTGGCGTGGCGCGTATCTCGATCGCGACGGCATGCCGAT 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 AATCCCGCGATTGAGTAGTATTCATGCTTGATCAGCGTGGCGCGCGACATGA 596
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 TCGCCGCGCGCGTGGTTACCAATCGCTGGTGCAACTGGTCCGCGCGCGTATCACGA 599
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 18, 2004, 23:07:12
Job time : 660.306 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 22:46:01 ; Search time 4048.28 Seconds
(without alignments)
8371.204 Million cell updates/sec

Title: US-10-759-889-1
Perfect score: 930
Sequence: 1 atggaacgactaaacgcatt.....ggcgagggaagaagaagtaa 930

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_gsa1:*
8: gb_gsa2:*
9: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	459	49.4	838	5	BQ144074
C 2	448	48.2	811	5	BQ143765
C 3	440	47.3	452	4	BG451124
C 4	187.8	20.2	205	4	BI406127
C 5	167.6	18.0	182	4	BI406112
C 6	136.8	14.7	165	4	BI406111
C 7	131.8	14.2	755	8	AF094919
C 8	105.8	11.4	703	9	BZ341676
C 9	101.4	10.9	565	9	CG708961
C 10	98.2	10.6	1027	9	CNS01FUZ
C 11	95	10.2	1178	8	BZ5533247
C 12	90.4	9.7	761	9	CG9899406
C 13	89	9.6	954	8	BZ558369
C 14	87.6	9.4	1279	8	BZ567908
C 15	87.2	9.4	741	9	CG899239
C 16	84	9.0	1169	8	BZ579305
C 17	82.8	8.9	550	8	BH389688
C 18	81.2	8.7	626	7	C0540363
C 19	80.4	8.6	777	7	CNS823930
C 20	77	8.3	632	7	C0535994
C 21	76.6	8.0	1287	8	BZ573075
C 22	74.8	8.0	716	7	CNS822664
C 23	74.8	8.0	886	7	CK408871
C 24	74	8.0	531	7	CF887195

C 25	74	8.0	548	6	CB852476
C 26	74	8.0	548	7	CF891359
C 27	73.4	7.9	1241	8	BZ567267
C 28	72.4	7.8	795	9	CL694387
C 29	68.6	7.4	1021	8	BZ554085
C 30	67.8	7.3	789	7	CNS824816
C 31	67.2	7.2	566	4	CK329341
C 32	67.2	7.2	585	4	BG063353
C 33	67.2	7.2	585	7	CK334119
C 34	67.2	7.2	828	9	CL659241
C 35	66.8	7.2	866	8	BZ678395
C 36	65.4	7.0	1241	8	BZ559127
C 37	65.2	7.0	815	7	CNS82455
C 38	65.2	7.0	820	7	CO202387
C 39	65.2	7.0	1227	8	BZ577725
C 40	65	7.0	964	8	BZ558727
C 41	64	6.9	1123	8	BZ564676
C 42	61.6	6.6	730	5	BM964257
C 43	61.4	6.6	1208	8	BZ552364
C 44	61.2	6.6	948	8	BZ578978
C 45	60.6	6.5	806	8	BZ551243

ALIGNMENTS

RESULT 1
BQ144074/c
LOCUS BQ144074 838 bp mRNA linear EST 24-APR-2002
DEFINITION NF032D03DT1F1027 Drought Medicago truncatula cDNA clone NF032D03DT
5', mRNA sequence.
ACCESSION BQ144074
VERSION BQ144074.1 GI:20281133
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; -
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 838)
Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.,
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula drought library
Unpublished (2000)
CONTACT: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6892
Email: gdmay@noble.org
Insert Length: 838 Std Error: 0.00
Plate: 032 row: D column: 03
Seq primer: TCACACAGGAACAGCATATCAC.
FEATURES
Location/Qualifiers
source
1..838
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF032D03DT"
/issue_type="plantlets"
/dev_stage="pooled timepoints"
/clone_lib="Drought"
/note="Vector: Lambda Zap; Contains a mixture of entire
plantlets harvested in a series of days-post-watering
timepoints."

ORIGIN

Query Match 49.4%; Score 459; DB 5; Length 838;
Best Local Similarity 100.0%; Fred. No. 5.8e-133;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
source
1. .452
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF099A07DT"
/tissue_type="Plantlets"
/dev stage="Pooled timepoints"
/clone_lib="Drought"
/notes="Vector: Lambda Zap; Contains a mixture of entire
plantlets harvested in a series of days-post-watering
timepoints."

ORIGIN
Query Match 47.3%; Score 440; DB 4; Length 452;
Best Local Similarity 99.8%; Pred. No. 4.9e-127;
Matches 451; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 478 GCATGCCAATGTTGGTGGCGCGGAAAGCTATCTCACACAATACGGCAATACCGGAA 537
Db 452 GCATGCCAATGTTGGTGGCGCGGAAAGCTATCTCACACAATACGGCAATACCGGAA 393
QY 538 AAACCCGCGGATTTGAGTAGTCATTCATGCTTGTATACAGGTGGCGCCGACAAATGAA 597
Db 392 AAACCCGCGGATTTGAGTAGTCATTCATGCTTGTATACAGGTGGCGCCGACAAATGAA 333
QY 598 TTGAACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTTCGTG 657
Db 332 TTGAACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTTCGTG 273
QY 658 ACTAATGATCCGATGACGCTGTGGCTGCTGACGGGGGTGCGGGATCGCTACGTG 717
Db 272 ACTAATGATCCGATGACGCTGTGGCTGCTGACGGGGGTGCGGGATCGCTACGTG 213
QY 718 CCGCTGATGCTGGTGTATCAACAGATCAATCGTGGGAGCTGGAGATCTCTGTCGCGGT 777
Db 212 CCGCTGATGCTGGTGTATCAACAGATCAATCGTGGGAGCTGGAGATCTCTGTCGCGGT 153
QY 778 TACAGTCAGATCCAGCCCGGTTTATCGTTATATACCGAAAAAGATAAGTCGCGGTG 837
Db 152 TACAGTCAGATCCAGCCCGGTTTATCGTTATATACCGAAAAAGATAAGTCGCGGTG 93
QY 838 AAGGTACAGGTC-GTGATCACTCGCTGACGATTTATTTGTTGAGTCGGTAAATGTT 896
Db 92 AAGGTACAGGTCGTTGATCACTCGCTGACGATTTATTTGTTGAGTCGGTAAATGTT 33
QY 897 TCAGGATGACCGGCGCGGAAAGAGAAGT 928
Db 32 TCAGGATGACCGGCGCGGAAAGAGAAGT 1

RESULT 4
BI406127/c
LOCUS
DEFINITION
155C0 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
ACCESSION
BI406127
VERSION
BI406127.1 GI:15185541
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 205)
CROOKSHANKS, M., EMMERSEN, J., WELINDER, K.G. and NIELSEN, K.L.
The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL
FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE
21475600
PUBMED
11591384
COMMENT
Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 182
POLYA-No.
Location/Qualifiers
1. .182
/organism="Solanum tuberosum"
/mol_type="mRNA"

source
1. .205
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="Tuber"
/clone_lib="Mature tuber lambda ZAP"
/notes="Vector: Lambda ZAP"

FEATURES
source
1. .205
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="Tuber"
/clone_lib="Mature tuber lambda ZAP"
/notes="Vector: Lambda ZAP"

ORIGIN
Query Match 20.2%; Score 187.8; DB 4; Length 205;
Best Local Similarity 99.0%; Pred. No. 1.4e-47;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 198 CCAGGGTGGCGTGGTATGTTTCATGAAGTGCAGGATGTTTCATGAGCAACTGTATGCTT 257
Db 205 CCAGGGTGGCGTGGTATGTTTCATGAAGTGCAGGATGTTTCATGAGCAACTGTATGCTT 146
QY 258 CAATAACACCCCATCGGACGCTACGCATGCTGCTTCACTATGCGCAAAATGT 317
Db 145 CAATAACACCCCATCGGACGCTACGCATGCTGCTTCACTATGCGCAAAATGT 86
QY 318 TCTCGCGGGCTCACGCCAAATGCTGAAGAATACCCAGGTTTGAGCGTCAATCTGGT 377
Db 85 TCTCGCGGGCTCACGCCAAATGCTGAAGAATACCCAGGTTTGAGCGTCAATCTGGT 26
QY 378 TACCGGAATTC 388
Db 25 TACCGGAATTC 15

RESULT 5
BI406112/c
LOCUS
DEFINITION
155_1 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
ACCESSION
BI406112
VERSION
BI406112.1 GI:15185526
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 182)
CROOKSHANKS, M., EMMERSEN, J., WELINDER, K.G. and NIELSEN, K.L.
The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL
FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE
21475600
PUBMED
11591384
COMMENT
Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 182
POLYA-No.
Location/Qualifiers
1. .182
/organism="Solanum tuberosum"
/mol_type="mRNA"
```


/cultivar="Field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="Tuber"
/clone_lib="Mature tuber lambda ZAP"
/note="Vector: Lambda ZAP"

ORIGIN

Query Match 18.0%; Score 167.6; DB 4; Length 182;
Best Local Similarity 95.1%; Pred. No. 3.2e-41; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 9;

QY 199 CAGGGCTGCGTCGTATGTTTCATGAAGTGCAGGATGTTTCATGACCAACTGTATGCCCTTC 258
DB 182 CAGGGCTGCGTCGTATGTTTCATGAAGTGCAGGATGTTTCATGACCAACTGTATGCCCTTC 123

QY 259 AATAACACCCCATCGGACGCTAGCATGCTGCTTTTCAACTATGCGCAAAATGTT 318
DB 122 AATAACACCCCATCGGACGCTAGCATGCTGCTTTTCAACTATGCGCAAAATGTT 63

QY 319 CTCGCCGGGCTGACAGCCAAATGCTGAAGATACCCAGGTTTGAGCGTCAATCTGTT 378
DB 62 CTCGCCGGGCTGACAGCCAAATGCTGAAGATACCCAGGTTTGAGCGTCAATCTGTT 3

QY 379 AC 380
DB 2 AC 1

RESULT 6

BI406111/c
LOCUS 165 bp mRNA linear EST 11-SEP-2002
DEFINITION 155.0 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
sequence.

ACCESSION BI406111
VERSION BI406111.1 GI:15185525
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 165)
Crockshanks, K., Emmersen, J., Welinder, K.G. and Nielsen, K.L.
The potato tuber transcriptome: analysis of 6077 expressed sequence
tags

JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE 21475600
PubMed 11591384

COMMENT Contact: Karen G. Welinder
Institute for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 165
POLYA=No.

FEATURES
source

1..165
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="Tuber"
/clone_lib="Mature tuber lambda ZAP"
/note="Vector: Lambda ZAP"

ORIGIN

Query Match 14.7%; Score 136.8; DB 4; Length 165;
Best Local Similarity 95.0%; Pred. No. 1.6e-31; Indels 1; Gaps 1;
Matches 152; Conservative 0; Mismatches 7;

QY 236 TTCATGAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCATGCTGTT 295
DB 165 TTCATGAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCATGCTGTT 106

QY 296 CTTCAACTATGGCACAATAATGTTTCGCCGGGCTGACAGCCAAATGCTGGAAGATACC 355
DB 105 CTTCAACTATGGCAC-AAATGTTTCGCCGGGCTGACAGCCAAATGCTGGAAGATACC 47

QY 356 CAGGTTTGAGCGTCAATCTGGTTACCGGAATTCAGCCCC 395
DB 46 CAGGTTTGAGCGTCAATCTGGTTACCGGAATTCAGCCCATCC 7

RESULT 7

AF094919
LOCUS 755 bp DNA linear GSS 29-AUG-2000
DEFINITION AF094919 Salmonella typhimurium LT2, Lambda DASH II Salmonella
typhimurium genomic clone 906-T3, genomic survey sequence.

ACCESSION AF094919
VERSION AF094919.1 GI:4322761
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium

REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
1 (bases 1 to 755)
Wong, R.M.Y. and McClelland, M.
End Sequences of Salmonella typhimurium LT2 Lambda DASHII Clones,
Li-Cor
Unpublished (1999)
Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifesci.sdsu.edu
Class: shotgun.

FEATURES Location/Qualifiers
source 1..755
/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="906-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

ORIGIN

Query Match 14.2%; Score 131.8; DB 8; Length 755;
Best Local Similarity 84.6%; Pred. No. 9e-30; Indels 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 27;

QY 1 ATGGAACGACTAAACGACATGTCGGTGTTCGCAAGTAGTTGAATTCGCTCTTTTACC 60
DB 581 ATGGAACGACTAAACGACATGTCGGTGTTCGCAAGTAGTTGAATTCGCTCTTTTACC 640

QY 61 GCCGCCGCCAGACAGCTTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 120
DB 641 GCCGCCGCCAGACAGCTTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATGCGGAACTG 700

QY 121 GAGATGAGTTGCGAGTAAGCTGTTAAACCGTAGCAGCAGGAGCTTGGCTGA 175
DB 701 GAGATGAGTTGCGAGTAAGCTGTTAAACCGTAGCAGCAGGAGCTTGGCTGA 755

RESULT 8

BZ341676
LOCUS 703 bp DNA linear GSS 06-NOV-2002
DEFINITION BZ341676.1 WGS-Sbicolorf (JMI.07 adapted methyl filtered) Sorghum
bicolor genomic clone ic78a09 5', genomic survey sequence.

ACCESSION BZ341676
VERSION BZ341676.1 GI:24741474
KEYWORDS GSS.

```

clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 565)
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1119011 row: 33
Class: transposon-tagged.
Location/Qualifiers
1. 565
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/Al88/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1119 - RescueMu Grid A8"
/note="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 Kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid A8 was grown at UC San Diego in 2002. DNA
was extracted from leaf strips, double digested using
BamHI and BglII, and ligated to form circular plasmids.
DH10B cells were transformed and then screened on LB
plates with ampicillin."

```

Sorghum bicolor (sorghum)
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 703)
Rabinowicz,P.D., O'Shaughnessy A.L., Baija,V., Dedhia,N.,
Kazemburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
Zutaverni,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from sorghum bicolor (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ic78 row: a column: 09
Seq primer: -2LM13UnivRev
Class: shotgun
High quality sequence stop: 703.
Location/Qualifiers
1..703
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/clone="ic78a09"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-SbcolorF (JM107 adapted methyl filtered)"
/note="Site 1: Xba I; Site 2: Xba I; The vector was
digested with XbaI and one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (.x/y reads in M13mp19,
.b/g reads in pUC19). The same ligation was transformed in
either JM107 or DH5a."

Query Match	10.9%	Score 101.4	DB 9	Length 565
Best Local Similarity	50.5%	Mid. No. 3.2e-20		
Matches 274	Conservative 0	Predictions 266	Indels 3	Gaps 1
QY	10	CTAAACGCGATGTCGGGTGTTTGCCAAAGTAGTTGAATTGGCTCTTTTACGCGCGCGCC	69	
DB	22	CTGAACGACATCGCGCATTCGTCGCTCGGTGGATACCGNAGTTTTTACGGCGGTGCC	81	
QY	70	AGACAGCTACAGATGAGCGGTTTGTCCTCATCTAGACGGGTATCAAAACTCGAAGATGAG	129	
DB	82	AGGCAACTCGGGGTGACCCCGCTCGGCGGTGGTAAATCGATCGTCGGCTTGAATCGCGT	141	
QY	130	TTGCAGGTAAAGCTGTGTTAAACCGTAGCACACCGACGATTTGGCTTGACCGNAGCCGGTAGA	189	
DB	142	TTGCAGGTCCGCTGTCTCAACCGCACCCCGCGCTTGAGCTGACCGAGCATGGTCGG	201	
QY	190	ATTTACTACAGGGCTCGCGTCGTATGTCTCATGAAGTCGAGGATGTTTCATGAGCAACTG	249	
DB	202	GTGCTGTATGAGCGATGTGGGCGATCTTTCAGGATCTGGACGAGTCTCGAAGAGGCTTG	261	
QY	250	TATGCTTTCAATAACACCCCGCATCGGACGCTACCGATTTGCTCTTTCAACTATGSCA	309	
DB	262	CGGTTCCGCGTTCGACGCGCCCGCGCTTTCGCGATGAGCGCTGCGGATGACACTGGGC	321	
QY	310	CAAAATGTTCTCGCCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGTTTGAGCGTC	369	
DB	322	CGGCTGCACGTTGTGCAGCACATCGAGCGCTCGCTTGAGAGACTGSCCGTCGCTCGCCATC	381	
QY	370	AATCTGGTTACCGAATTCAGCCCCCGACCTGATTCGGACGGCTCTGGATGTCGTTGATC	429	
DB	382	GACGAGACGTTTTCCGATCGGCTGGTCGACCTGATCGACGAGGGTTTCGATCTCGCGATT	441	
QY	430	CGCGTCGGCGCG---TTGCAGGATTCACGCGCTGTTTTCCCGCGGCTCTGGGCGGATGCCA	486	

Query Match	11.4%;	Score 105.8;	DB 8;	Length 703;	
Best Local Similarity	72.5%;	Pred. No. 1.4e-21;			
Matches 137;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;	
QY	715	GTGCGCGTGATGTGGGTGATCAACAGAGATCAATCGTGGGGAGCTGGAGATCCTGCTCCG	774		
Ddb	1	GTGCGCGTGATGTGGGTGATGATGAGATTACGCCGGTGAGATCGATTTCCTCCG	60		
QY	775	CGTTACAGATCAGATCACGCCCGGTTTATGCGTTATATATACGAAAAAGATAAGCTGCCG	834		
Ddb	61	CAGTATCACTCGGAGCGCGCCCGGTGTATCGCGTCTATACCGAGAAGGATAAGCTGCCG	120		
QY	835	CTGAAGCTACAGGTGCGTGATCAACTCGCTGACGGATTATTTTGTGAGTCCGTAAATTG	894		
Ddb	121	CTCAGGTACAGGTGTGTATCAATTATCTGACCGAGTATTTACGCAGGTCCGACGGCAG	180		
QY	895	TTTCAGGAG 903			
Ddb	181	TATCAGCAG 189			
RESULT 9					
LOCUS	CG708961				
DEFINITION	ll19011C12.1EL.y1 ll119 - RescueMu Grid AA Zea mays genomic, genomic survey sequence.				
ACCESSION	CG708961				
VERSION	CG708961.1				
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
	CG708961 ll19011C12.1EL.y1 ll119 - RescueMu Grid AA Zea mays genomic, genomic				
	565 bp	DNA	linear	GSS 20-OCT-2003	

```

Db      442  CGCATCGTCCGCCAGGAAGACTCCGACTACTCAGCGGACGGTTCAGCAGAA 501
Qy      487  ATGGTGGTGGCGCGCGGAAAGCTATCTCACAAATACGGGATACCGGAAACCCGCC 546
Db      502  ATGATCACCTGTGCTCGCCGCGGATTATCTGAAGAAACATCCGCCACCGCAGACCCGAA 561
Qy      547  GAT 549
Db      562  GAT 564

RESULT 10
CNS01FUZ
LOCUS
DEFINITION
  Anopheles gambiae GSS T7 end of clone 05P23 of NotreDamel library
  from strain PEST of Anopheles gambiae (African malaria mosquito),
  genomic survey sequence.
ACCESSION
  AL142332
VERSION
  AL142332.1 GI:7000450
SOURCE
  GSS.
ORGANISM
  Anopheles gambiae (African malaria mosquito)
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
REFERENCE
  1 (bases 1 to 1027)
  Direct Submission
  Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  2 (bases 1 to 1027)
  Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
  Direct Submission
  Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
  Roux, Paris 75015, France
  This clone is from an A. gambiae BAC library provided by F.H.
  Collins and sequenced by Genoscope in collaboration with the
  Laboratory of Biochem. and Biol. Molec. of Insects, Institut
  Pasteur.
FEATURES
  source
  1..1027
  /organism="Anopheles gambiae"
  /mol_type="genomic DNA"
  /strain="PEST"
  /db_xref="taxon:7165"
  /clone="05F23"
  /clone_lib="NotreDamel"
  /note="end : T7"

ORIGIN
Query Match 10.6%; Score 98.2; DB 9; Length 1027;
Best Local Similarity 47.3%; Pred. No. 3.9e-19;
Matches 375; Conservative 5; Mismatches 408; Indels 5; Gaps 3;

Qy      76  CTACAGATGAGCGTTTCGTCATCGTCAGAGCGGTATCAAACTCGAAGATGAGTTGAG 135
Db      20  CTGCAATTTGTCAGCGTCGGCGGTAGAAAAGCATCGCTCGTCTCGAGAGCGACTGGGG 79

Qy      136  GTAAGCTGTTAAACCGTAGCAGCAGCATTTGCGCTTGACCGAAGCCGGTAGAATTTAC 195
Db      80  GTACGGCCATTTACGGCGCACCCCGCAGCCAGARCTTGACCGATACCGGCCCTTTT 139

Qy      196  TACAGGGCTGGCGTGGTATGTTATGATGAGTGCAGGATGTTTCATGAGCACTGTATGCC 255
Db      140  TATGAACGGTGGCTGGCTGGCTGGAGGAAATTCGGGGCGCAGAGTCGCTGTTGAAACG 199

Qy      256  TTCAATAACACCCCATCGGGAGCGGTACGCATTTGGCTGTCTCAACTATGGCAAAAAT 315
Db      200  GGGAAACAGCAGGTACAGCGCGCCGCTGCGCGTGTGCATGCGCGTACTGTTGGCGCCAG 259

Qy      316  GTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGTTTGAGCGTCAATCTG 375

```

```

Db      260  TGGTTCGCGCGCTGTTGATTAACTTGGCGCAGGAACACCCCGGGCTTARCTTGAATG 319
Qy      376  GTTACCGGAATTCACGCCCCGACCTGATTCGCGAGGCTCTGGATGTGGTATCCGCGTC 435
Db      320  TCGTTTACGCGACCGCGTGGTGGATCTGGTGGAGGAGGGTTTGAATATGGCGGTGGGTAAC 379
Qy      436  GCGCGCTTGCGAGATTCCAGCCCTGTTTTCCCGCCGCTCG--GCGCGCATGCCAAATGGTGG 493
Db      380  GGCACTGTCAGCAGCAGCAGCATGCTGTCGCGCAGAAARCTGGGGAAACACCCGATGTC 439
Qy      494  TGTGCGCGCGGAAAAGCTATCTCACAAATACGGGATACCGGAAAACCCGCGGATTTGA 553
Db      440  TGTGCGCTGCGCGGAGATTATCTRCAGAAACAAAGCCACCCGCAAAAGCGTGGTGACTTAC 499
Qy      554  GTAGTCATTATCGCTTTGAATA-CAGCGTGGCGCCGACAAATGAATTTG--AACTGATCG 610
Db      500  CCATATATAGGCCATTAATCTATCTGCGCGCAGGAGAGTGTACCTTGCAACTGATCG 559
Qy      611  CACCGGAAGGGATCTCGACTCGCTGATCCCAAGAAAGATTGTGACTAATGATCCGA 670
Db      560  ATAAAGGAAGGACCTCGCACACCTTTACTCCCGCTTCATCGCTCAATATGGATGATTTGC 619
Qy      671  TGAGCGTGGTGGCTGCTGACGGCGGGTCCGGGATCCCTACGTGCGCGCTGATGGG 730
Db      620  AGCCCATCTCGGACGCGCGCTGCGCGGACACGGGATTCCTGGCTCCCTCTGCGATGG 679
Qy      731  TGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGTCGCGGTTTACCACTGAGATC 790
Db      680  CCAATTAAGAAATTCATCAGGGAAGCTCTGTCGCTCTCTGAAACAGGCGCCGATGTC 739
Qy      791  CAGCGCGGTTTATGCTTATATACCGAAAAGATTAAGTTCGCGTGAAGGTACAGGTG 850
Db      740  ATTTGACGCTCATGCTGCTGTCGACGAAACCGCCACCTGCGCGTGGAGTGGAGATTG 799
Qy      851  TGATCAACTCGCT 863
Db      800  CCGTCGATAGCT 812

RESULT 11
BZ553247 1178 bp DNA linear GSS 17-DEC-2002
pacsl-60_405.sl pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_405, genomic survey sequence.
ACCESSION BZ553247
VERSION BZ553247.1 GI:27159019
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
  Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
  Pseudomonadaceae; Pseudomonas.
REFERENCE
  1 (bases 1 to 1178)
  Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
  Burns,J.L., Kaul,R. and Olser,M.V.
  Whole-Genome-Sequence variation among multiple isolates of
  Pseudomonas aeruginosa library
  J. Bacteriol. (2002) In press
  Contact: Chris K. Raymond
  Genome Center
  University of Washington
  Box 35245, Seattle, WA 98105-2145, USA
  Tel: 2062216954
  Fax: 2066857244
  Email: craymond@u.washington.edu
  Class: shotgun.
  Location/Qualifiers
  1..1178
  /organism="Pseudomonas aeruginosa"
  /mol_type="genomic DNA"
  /strain="1-60"
  /db_xref="taxon:287"
  /clone="pacsl-60_405"

```

```

/clone_lib="pacsl-60"
/note="clinical isolate 1-60 whole genomic shotgun
library."

ORIGIN
Query Match      10.2%; Score 95; DB 8; Length 1178;
Best Local Similarity 50.1%; Pred. No. 4.1e-18;
Matches 262; Conservative 0; Mismatches 260; Indels 1; Gaps 1;

QY 1 ATGGACGACTAAAGCATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 139 ATGGATTCGCTGAAGGCAATGCGATCTTCGCCACTGTGTGGACAAGGCTCGATGCA 198
QY 61 GCGCGCGCCAGACAGCTACAGATGAGCGTTTCCTCCATCAGTCAGACAGTATCAAACTG 120
DB 199 GCGCGCGCCAGACAGCTGCGCATGACCCCTTCGGCAGTCAGCAGCAGATCGCAAGCTG 258
QY 121 GAAGATGAGTTCAGGTAAGCTGTAAACCTAGCACACGAGCATTGGCTGACCGAA 180
DB 259 GAGAGCGCGCCAGGTCACTTTGTGTCATCGCACACCGCGCGCTGACCTGACCGAG 318
QY 181 GCGGCTAGAAATTTACTACAGCGGCTGCGGTTCGTATGCTTCATGAAGTCAGGATGTTTCA 240
DB 319 GCGCGCGAGCGTTCATCGCAGTTGCGCGAGATGCTGGCGATCGCCGAGGAGCCGAG 378
QY 241 GAGCAACT-GTATGCGTTCAATAACACCCCCATCGGAGCGCTACGCAATGGCTGTTC 299
DB 379 CGCGCGCTCGCGGTAAATGCGCGGATGCGCGGTGGCGGATGTCGCTGGCGCGCGGT 438
QY 300 AACTATGGCACAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAGAAATACCCAGG 359
DB 439 GGGGTTCTCGGACGCTGATCACCGAGCGCTGAAACGGTTCTGGAGAACCCCGCA 498
QY 360 TTGAGCGCTCAATCTGGTTACCGGAATTCAGCCCCCGACCTGATTCGCGAGCGTCTGA 419
DB 499 GTTGCGCTCGCACTGTTCTTCAGGACGAGGCGCATGATCTGTCGCGGAGCGCATCGA 558
QY 420 TGTGTGATCCGCTCGCGCTTCGAGATTCAGCGCTGTTTCCGCGCTCTGGGCGC 479
DB 559 CTGGCGATCCGCGTGGCAATCTCGCGACTCCAGCTGGTGGCGCGCCACCTCGGCGA 618
QY 480 GATGCCAATGTGTGTGCGCGCGCAAAAGCTATCTCACAA 522
DB 619 CTGAGCAGCGTGTCTGCGCGCGCGCGCTATTTGCGCAA 661

RESULT 12
CG899406/c
LOCUS
DEFINITION
pacsl-60_962.sl pacsl-60 Pasteuria penetrans genomic 5',
genomic survey sequence.
ACCESSION
CG899406
VERSION
CG899406.1 GI:39554967
KEYWORDS
GSS.
SOURCE
Pasteuria penetrans
ORGANISM
Pasteuria penetrans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Pasteuria.
REFERENCE
1 (bases 1 to 761)
Opperman,C.H., Davies,K.G., Sosinski,B.R., Waterman,J. and Burke,M.
Unpublished Data
Unpublished (2003)
Contact: Opperman CH
Center for the Biology of Nematode Parasitism and Nematode
Interactions Unit
North Carolina State University and Rothamsted Research, Ltd.
Box 7233, NCSU, Raleigh, NC 27606, USA
Email: warthog@unity.ncsu.edu
Homology: e-val = 1e-40. Description = PROBABLE TRANSCRIPTION
REGULATOR PROTEIN [Ralstonia solanacearum]
gi|17431583|emb|CAD18261.1| PROBABLE TRANSCRIPTION REGULATOR
PROTEIN [Ralstonia solanacearum] Homology: e-val = 1e-40.
Description = PROBABLE TRANSCRIPTION REGULATOR PROTEIN [Ralstonia
solanacearum] gi|17428746|emb|CAD15431.1| PROBABLE TRANSCRIPTION

```

```

REGULATOR PROTEIN [Ralstonia solanacearum] row: f column: 7
Class: shotgun
High quality sequence stop: 761.
Location/Qualifiers
1. 761
/organism="Pasteuria penetrans"
/mol_type="genomic DNA"
/strain="Res147"
/db_xref="taxon:86005"
/clone_lib="Res147.1"

ORIGIN
Query Match      9.7%; Score 90.4; DB 9; Length 761;
Best Local Similarity 49.3%; Pred. No. 1e-16;
Matches 266; Conservative 0; Mismatches 271; Indels 3; Gaps 1;

QY 10 CTAAACGCAATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC GCGCGCC 69
DB 573 CTGGAATTTGATGATATTTTGTTCAGTCGTGTGAATTCGTTAGTTTACCGCGCGCT 514
QY 70 AGACAGCTACAGATGAGCGTTCGTCATCAGTCAGACGGTATCAAACTTGAAGATGAG 129
DB 513 GAGCGCTTACAGCTGCGATCGCCCTGCGGTGAGTTCGCTACACAGCTTGAAGACAAG 454
QY 130 TTGAGGTAAGCTGTAAACCTAGCACAGCAGCATTTGGCTGACCGAAGCGGTAGA 189
DB 453 CTTGGCGTTAAGCTATTGACCGCACCGCAGTCTGAGTATGACCGTTGAAGGGGAA 394
QY 190 ATTACTACAGCGCTGCGCTGCTGCTCATGTAAGTGCGAGTGTTCATGAGCAACTG 249
DB 393 GAAATTTATCAGCGGCTAATAACGCTACTGAGTGTGCGACGATAT---CATGGCTTCT 337
QY 250 TATGCGCTTCAATAACACCCCATCGGAGCGCTACGAGTGGCTGTCTTCAACTATGGCA 309
DB 336 TACTCTTCAATCAGGACCTCGTGGCCAGCTCGCATTTGATTTCCACTGGCGCTGGCA 277
QY 310 CAAAATGTTCTCGCGGCTGACAGCCAAATCTGAAAGATACCCAGTTTGAGCGTC 369
DB 276 CACAGTATCTGATCCCAATTAAGCGAGTTTAAGGTTTATCTGAAATTTGAATA 217
QY 370 AATCTGTTTACCAGAAATCCAGCCCCCGACCTGATTCGCGACGGTCTGGATGGTGATC 429
DB 216 GTCTTAACCGCTTCAGATCGCTGACCCATTTAATTCGAGAAGGTGTGACTGTGTATT 157
QY 430 CGGTCGCGCGGTGAGGATTCAGCGCTTTTCCGCGCTGTCGGCGGATGCCAATG 489
DB 156 CGCCTCGGTGAGCTTGTGATTCGAATTTTATTCGCGCAGAAATGGCAGCTGCGAGTG 97
QY 490 GTGCTGTGCGCGCGCAAAAGCTATCTCACAAATAGCGCATACCGGAAACCGCGCGAT 549
DB 96 GTACCTGTGACGACACCTCTTATCTGGATAAATATGGCACTCTCTTAACGCTGATGAT 37

```

```

RESULT 13
BZ558369
LOCUS
DEFINITION
pacsl-60_962.sl pacsl-60 Pasteumonas aeruginosa genomic clone
pacsl-60_962, genomic survey sequence.
ACCESSION
BZ558369
VERSION
BZ558369.1 GI:27172622
KEYWORDS
GSS.
SOURCE
Pasteumonas aeruginosa
ORGANISM
Pasteumonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteumonadales;
Pasteumonadaceae; Pasteumonas.
REFERENCE
1 (bases 1 to 954)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pasteumonas aeruginosa library
J. Bacteriol. (2002) in press
Contact: Chris K. Raymond
Genome Center

```

University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .954
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacs1-60.962"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

Query Match 9.6%; Score 89; DB 8; Length 954;
Best Local Similarity 49.2%; Pred. No. 38-16; Indels 2; Gaps 1;
Matches 283; Conservative 0; Mismatches 270; Indels 2; Gaps 1;

QY 2 TGAACGACTAAACGCGATGCGGTGTTTGCACAAAGTAGTTGAATTTGGCTCTTTTACCG 61
DB 75 TGGATCACCTACCGCTCTCAAGGTCTTCGCGCGCTGCGCGAATGGCGCTTCGCG 134
QY 62 CGCCGCCAGACAGCTACAGATGAGGTTTCGTCCATCAGTCAGTCAGCGGTATCAAACTGG 121
DB 135 CGCGCGCCGCGAGATGAATCTCTCGCGCGCGCGCTGAGCAAGAACGTCCCGGAGCTGG 194
QY 122 AAGATGAGTTGAGGTAAGAGCTGTTAAACCGTAGCACAGCAGCATTTGGCTTCAGCCGAAG 181
DB 195 AAGCGACCTCAAGGTGCCCATGATCAATCGCACCCCGCAGCATGAGCCTGACCGAGG 254
QY 182 CGGTGAGATTTACTACGAGGCTGCGCTGATGCTTCATGAGTGCAGATGTTCTATG 241
DB 255 CGCGGGAAGTCTACCGGAGCGCTTGGAGCGCATCTCGACGACCTCGAGCGCCCGGACG 314
QY 242 AGCAACTGATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTTGGCTTCTTCAA 301
DB 315 CGCGCTCACTTGATGAGAGGCGCCCGAGCGCTGCTGGGCTCAGCGCCGCTGA 374
QY 302 CTATGGCAAAATGTTCTGCGCGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGTT 361
DB 375 CCTCGCCCTACCTGCTGACCCCGGCATTCGCGCTTTTCTCCAGCGTTATCCGAGC 434
QY 362 TGAGGCTCAATCTGTTACCGAATTCAGCCCGCGAGCTGATTCGCCAGCTCTG--GA 419
DB 435 TCGCCCTGGAATGCTCTCTGAGAGCGCGCGCAGACCTGATCGCCGAAGATCGACCT 494
QY 420 TGTGTGATCCCGCTGCGCGCTTGAGGATTCAGACCTGTTTCCCGCGCTCTGGGCGC 479
DB 495 GGCCTGCGTGTAGCGACCGGTTGCGGACTCCGCGCTGCTGGCGCGCGCTGCTGT 554
QY 480 GATGCCAATGTTGTTGCGCGCGGAAAGCTATCTCACAAATACGACATACCG 534
DB 555 CTTGGAGCAGCTGCTCTGCGCGCGCGCGCTTACCTCAGTCAGCATGCGCCGCG 609

RESULT 14

BZ567908 1279 bp DNA linear GSS 17-DEC-2002
LOCUS pacs2-164_7276.xl pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION pacs2-164_7276, Genomic survey sequence.
ACCESSION BZ567908
VERSION BZ567908.1 GI:27200960
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1279)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

TITLE Whole-Genome-Sequence variation among multiple isolates of
JOURNAL Pseudomonas aeruginosa library
COMMENT J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .1279
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_7276"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 9.4%; Score 87.6; DB 8; Length 1279;
Best Local Similarity 49.0%; Pred. No. 9e-16; Indels 0; Gaps 0;
Matches 234; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCGATGCGGTGTTTGCACAAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 114 ATGGAATTCGCTGAAGGCGATGGCGATCTTCGCGACCTGAGGTGGAACAAGGCTCGATGGCA 173
QY 61 GCGCGCGCAGACAGCTACAGATGAGCGTTTTCGTCCATCAGTCAGACGGTATCAAAACTG 120
DB 174 GCGCGCGCAGAGCTTGGCGATGACCCCTTGGCAGTCAGCCAGCATCCGCAAGCTG 233
QY 121 GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACAGCAGCATTTGGCTGACCGAA 180
DB 234 GAGAGCGCGCGCAGGTCATCTTGTGTCATCGCACCCCGCGGCTGACCCCTGACCGAG 293
QY 181 GCGGTGAGATTTACTACAGGCTGCGCTGCTGATGCTTCATGAGTGCAGATGTTTCA 240
DB 294 GCGCGCGAGCGTTCTATCGAGTTGCGCGAGATCTGGCGATCGCCAGAGCCGAG 353
QY 241 GAGCAACTGTTATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTTGGCTTCTTCA 300
DB 354 CGGTGCTTGGCGAATGGCGCATGCGCGGTGGGTGAGTTGCGCTGCGCGCACCGAG 413
QY 301 ACTATGGCAAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAGAAATACCCAGT 360
DB 414 GGTCTTCCGCGACGCTGATCACCCAGGCGCTGAATTCGTTTGTGAGAACCCAGCAG 473
QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCGCCCGACCTGATTCGCCAGCGTCTGGAT 420
DB 474 TTGCGCTCAGCTTGTCTTCCAGGAGCGCATCGAGCTGGGCGCGGCGCATCGCA 533
QY 421 GTGGTATCGCGCTGCGCGCTTTCAGGATTCAGGCTGTTTTCGCGCGCTGCGCGC 478
DB 534 CTGACGATACGCGCTCGGACATTTCTCCGACTCAGCCTGTGTCGACGCACTTTTGTG 591

RESULT 15
CG89239/c
LOCUS pastbac094xd07.bl.ab1 Res147 1 Pasteuria penetrans genomic 5',
DEFINITION genomic survey sequence.
ACCESSION CG89239
VERSION CG89239.1 GI:39554734
KEYWORDS GSS.
SOURCE Pasteuria penetrans
ORGANISM Pasteuria penetrans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Pasteuria.
REFERENCE 1 (bases 1 to 741)

Search completed: December 19, 2004, 02:46:57
Job time : 4053.28 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:49:04 ; Search time 1924.01 Seconds
(without alignments)
9659.452 Million cell updates/sec

Title: US-10-759-889-3

Perfect score: 393
Sequence: 1 tccaactcaaacgaacagt.....ctttaagtcagagtgaata 393

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hhg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	393	100.0	110000	1	ECOUW67_1
2	393	100.0	110000	1	U00096_33
3	393	100.0	290452	2	AC079167 Mus muscu
4	391.4	99.6	292497	1	AE016989 Shigella
5	389.8	99.2	307962	1	AP002564 Escherich
6	386.6	98.4	300359	1	AE016767 Escherich
7	362.8	92.3	115904	1	AE005551 Escherich
8	289.8	73.7	21252	1	AE008855 Salmonell
9	289.8	73.7	258050	1	AL627278 Salmonell
10	289.8	73.7	299991	1	AE016845 Salmonell
11	226	57.5	11744	1	AE015338 Shigella
12	225.4	57.4	11734	1	AE015339 Shigella
13	146.6	37.3	258	6	AR389738 Sequence
14	125.8	32.0	593	1	AB080601 Serratia
15	117.4	29.9	10553	1	AE013617 Yersinia
16	117.4	29.9	229338	1	AE017142 Yersinia
17	117.4	29.9	235050	1	AJ414158 Yersinia
18	100.4	25.5	110000	1	ECOUW67_1
19	100	25.4	100	6	AX988997 Sequence

C 20	100	25.4	100	6	AX988998	Sequence
C 21	100	25.4	100	6	AX988999	Sequence
C 22	87	22.1	11593	1	AE005552	Escherich
C 23	65.8	16.7	20157	1	AY647257	Yersinia
C 24	45.6	11.6	302614	1	AE016778	Pseudomon
C 25	44.4	11.3	1512	1	AB077386	Serratia
C 26	44.2	11.2	7218	6	I66494	Sequence 14
C 27	42.4	10.8	348525	1	AX640428	Bordetell
C 28	41.8	10.6	171655	2	AX950210	Danio rer
C 29	41.8	10.6	193489	5	AX005088	Zebrafish
C 30	41.8	10.6	217867	2	EX323049	Danio rer
C 31	40.8	10.4	348642	1	EX640446	Bordetell
C 32	40.6	10.3	303226	1	AE016774	Pseudomon
C 33	40	10.2	154091	2	CR394557	Danio rer
C 34	38.8	9.9	349174	1	AB063522	Wiggleswo
C 35	38.4	9.8	986	8	AB118012	Oryza sat
C 36	38.4	9.8	1181	6	E14788	CDNA encodi
C 37	38.4	9.8	157205	8	AP005098	Oryza sat
C 38	38.2	9.7	109619	2	AC141815_3	Continuation (4 of
C 39	37.8	9.6	85234	8	AC139743	Medicago
C 40	37.8	9.6	224562	10	AC123880	Mus muscu
C 41	37.6	9.6	205587	2	AC139761	Mus muscu
C 42	37.2	9.5	1035	6	AR384449	Sequence
C 43	37.2	9.5	50434	2	AC115729	Mus muscu
C 44	37.2	9.5	181337	10	AL929046	Mouse DNA
C 45	37.2	9.5	189589	2	AC146641	Otolemur

ALIGNMENTS

RESULT 1

ECOUW67_1

WPCOMMENT

Sequence split into 4 fragments LOCUS ECOUW67 Accession U18997

Fragment Name Begin End

ECOUW67_0 1 110000

ECOUW67_1 100001 218000

ECOUW67_2 200001 310000

ECOUW67_3 300001 372438

Continuation (2 of 4) of ECOUW67 from base 100001 (U18997 Escherichia coli K-12 chromoso

Query Match

Best Local Similarity 100.0%; Score 393; DB 1; Length 110000;

Mismatches 393; Conservative 0; Indels 0; Gaps 0;

Qy 1 TTCAACTCAAAAGCAAGTCGGGATATCAATAAAGCAATTCGTTGGGAGGACAGCC 180

Db 69872 TTCAACTCAAAAGCAAGTCGGGATATCAATAAAGCAATTCGTTGGGAGGACAGCC 70051

Qy 61 TGAACACGCGGATGCCAGAACGTCGTAGATACCTGTTGGCACAAGTACCGCGGCA 120

Db 69932 TGAACACGCGGATGCCAGAACGTCGTAGATACCTGTTGGCACAAGTACCGCGGCA 69991

Qy 121 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCGTTGGGAGGACAGCC 180

Db 69992 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCGTTGGGAGGACAGCC 70051

Qy 181 CAAACACCAAGTAAACGGGAAACAGACTCATGTGTGACCTTGGTTGTAAGAGAGAGCAGG 240

Db 70052 CAAACACCAAGTAAACGGGAAACAGACTCATGTGTGACCTTGGTTGTAAGAGAGAGCAGG 70111

Qy 241 CGTTATTATTTTCAGCATCTGTCGCGCAGAGAAGGCGATGGAAGCCGGCGAGAGCAA 300

Db 70112 CGTTATTATTTTCAGCATCTGTCGCGCAGAGAAGGCGATGGAAGCCGGCGAGAGCAA 70171

Qy 301 CATTGCTGTAGATGATATTTAATATATATTAGCGTAACCTGTTATGCTGTTATCTATATTAT 360

Db 70172 CATTGCTGTAGATGATATTTAATATATATTAGCGTAACCTGTTATGCTGTTATCTATATTAT 70231

Qy 361 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 393

Db 70232 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 70264

RESULT 2
U00096_33
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (34 of 47) of U00096 from base 3300001 U00096 Escherichia coli K-12 MG1655

Query Match 100.0%; Score 393; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8e-104;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCACTCAAGCAAGTCGCGATATCAATAAAGAGAGCAAGCAAGTACGCGGGTGT 60
Db	87149	TTCACTCAAGCAAGTCGCGATATCAATAAAGAGAGCAAGTACGCGGGTGT 87208
Qy	61	TGAACAACGCGGATCCGACAAAGTCGTAGTACTCTTGGCAAGTACCCGCGCA 120
Db	87209	TGAACAACGCGGATCCGACAAAGTCGTAGTACTCTTGGCAAGTACCCGCGCA 87268
Qy	121	CCAGCCAGAAATCGCGATGATAAAGCAATTCAAAAATATCGTGGAGGACGCC 180
Db	87269	CCAGCCAGAAATCGCGATGATAAAGCAATTCAAAAATATCGTGGAGGACGCC 87328
Qy	181	CAACACCCAGATAACCGGAAACAGACTCATGTGTGACCTTGTGTTAAAGAGAGCAGG 240

Db	87329	CAACACCCAGATAACCGGAAACAGACTCATGTGTGACCTTGTGTTAAAGAGAGCAGG 87388
Qy	241	CGTTATTATTTTCAGCATCTGTGCGCAGAGAGGCGATGAAAGCCGGCGAGAGCAA 300
Db	87389	CGTTATTATTTTCAGCATCTGTGCGCAGAGAGGCGATGAAAGCCGGCGAGAGCAA 87448
Qy	301	CATTGCTGTAGATTGATATTAATATATTAGCGTACTCTTATGCTGTTATCTATATTAT 360
Db	87449	CATTGCTGTAGATTGATATTAATATATTAGCGTACTCTTATGCTGTTATCTATATTAT 87508
Qy	361	GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
Db	87509	GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 87541

RESULT 3
AC079167/c

LOCUS
DEFINITION
AC079167 290452 bp DNA linear HTG 15-MAY-2002
SEQUENCING IN PROGRESS *** 196 unordered pieces.

ACCESSION
AC079167.2 GI:9945047

VERSION
HTG; HTGS PHASE1.

KEYWORDS
Mus musculus (house mouse)

SOURCE
Mus musculus

ORGANISM
[Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 290452)

REFERENCE
AUTHORS
Grillis, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.

TITLE
High Throughput Mouse Sequencing

JOURNAL
Unpublished

REFERENCE
2 (bases 1 to 290452)

AUTHORS
Grillis, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.

TITLE
Direct Submission

JOURNAL
Submitted (23-AUG-2000) Department of Molecular Genetics, Albert

Einsteine College of Medicine Genome Center, 1300 Morris Park Ave.,

Bronx, NY 10463, USA

On Aug 30, 2000, this sequence version replaced gi:9885883.

COMMENT

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpcgg.org/Sequence/mouse.html>

Contact: hpcg@mcg.harvard.edu

-----Summary Statistics

Center project name: AEN

Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 196511 at least Q20

*Consensus quality: 172308 at least Q30

*Consensus quality: 138790 at least Q40

Estimated insert size: agarose-FP - N/A

**Estimated coverage: agarose-FP - N/A

Quality coverage: 3.3 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 196 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 852: contig of 852 bp in length

* 833 872: gap of unknown length

* 873 1753: contig of 881 bp in length

* 1754 1773: gap of unknown length

* 1774 2658: contig of 885 bp in length

* 2659 2678: gap of unknown length
* 2679 3514: contig of 836 bp in length
* 3515 3534: gap of unknown length
* 3535 4449: contig of 915 bp in length
* 4450 5384: gap of unknown length
* 5385 5404: contig of 915 bp in length
* 5405 6273: gap of unknown length
* 6274 6293: contig of 869 bp in length
* 6294 7035: gap of unknown length
* 7036 7115: contig of 802 bp in length
* 7116 7981: gap of unknown length
* 7982 8001: contig of 866 bp in length
* 8002 8729: gap of unknown length
* 8730 8749: contig of 728 bp in length
* 8750 9875: gap of unknown length
* 9876 9895: contig of 1126 bp in length
* 9896 10631: gap of unknown length
* 10632 10651: contig of 736 bp in length
* 10652 11577: gap of unknown length
* 11578 11597: contig of 926 bp in length
* 11598 12454: gap of unknown length
* 12455 12474: contig of 857 bp in length
* 12475 13334: gap of unknown length
* 13335 13354: contig of 860 bp in length
* 13355 14279: gap of unknown length
* 14280 15245: contig of 925 bp in length
* 15246 15265: gap of unknown length
* 15266 16058: contig of 793 bp in length
* 16059 16078: gap of unknown length
* 16079 16945: contig of 867 bp in length
* 16946 16965: gap of unknown length
* 16966 17850: contig of 885 bp in length
* 17851 17870: gap of unknown length
* 17871 18722: contig of 852 bp in length
* 18723 18742: gap of unknown length
* 18743 19556: contig of 814 bp in length
* 19557 19576: gap of unknown length
* 19577 20816: contig of 1240 bp in length
* 20817 20836: gap of unknown length
* 20837 21756: contig of 920 bp in length
* 21757 21776: gap of unknown length
* 21777 22517: contig of 741 bp in length
* 22518 22537: gap of unknown length
* 22538 23457: contig of 920 bp in length
* 23458 23477: gap of unknown length
* 23478 24289: contig of 812 bp in length
* 24290 24309: gap of unknown length
* 24310 25249: contig of 940 bp in length
* 25250 25689: gap of unknown length
* 25690 26135: contig of 866 bp in length
* 26136 26155: gap of unknown length
* 26156 27048: contig of 893 bp in length
* 27049 27068: gap of unknown length
* 27069 28003: contig of 935 bp in length
* 28004 28023: gap of unknown length
* 28024 28874: contig of 851 bp in length
* 28875 28894: gap of unknown length
* 28895 29777: contig of 883 bp in length
* 29778 29797: gap of unknown length
* 29798 30160: contig of 363 bp in length
* 30161 30180: gap of unknown length
* 30181 31014: contig of 834 bp in length
* 31015 31034: gap of unknown length
* 31035 31880: contig of 846 bp in length
* 31881 31900: gap of unknown length
* 31901 33160: contig of 1260 bp in length
* 33161 33180: gap of unknown length
* 33181 34033: contig of 853 bp in length
* 34034 34053: gap of unknown length
* 34054 34972: contig of 919 bp in length
* 34973 34992: gap of unknown length
* 34993 35875: contig of 883 bp in length
* 35876 35895: gap of unknown length
* 35896 36731: contig of 836 bp in length
* 36732 36751: gap of unknown length
* 36752 37639: contig of 888 bp in length
* 37640 37659: gap of unknown length
* 37660 38499: contig of 840 bp in length
* 38500 38519: gap of unknown length
* 38520 39348: contig of 829 bp in length
* 39349 39368: gap of unknown length
* 39369 40377: contig of 1009 bp in length
* 40378 40397: gap of unknown length
* 40398 41469: contig of 1072 bp in length
* 41470 41489: gap of unknown length
* 41490 41891: contig of 401 bp in length
* 41891 41911: gap of unknown length
* 41912 42795: contig of 885 bp in length
* 42796 42815: gap of unknown length
* 42816 43744: contig of 929 bp in length
* 43745 43764: gap of unknown length
* 43765 44628: contig of 864 bp in length
* 44629 44648: gap of unknown length
* 44649 45958: contig of 1310 bp in length
* 45959 45978: gap of unknown length
* 45979 46827: contig of 849 bp in length
* 46828 46847: gap of unknown length
* 46848 47807: contig of 860 bp in length
* 47808 47827: gap of unknown length
* 47828 48376: contig of 549 bp in length
* 48377 48396: gap of unknown length
* 48397 49297: contig of 901 bp in length
* 49298 49317: gap of unknown length
* 49318 50182: contig of 865 bp in length
* 50183 50202: gap of unknown length
* 50203 51054: contig of 852 bp in length
* 51055 51074: gap of unknown length
* 51075 51373: contig of 299 bp in length
* 51374 51393: gap of unknown length
* 51394 52053: contig of 660 bp in length
* 52054 52073: gap of unknown length
* 52074 52944: contig of 871 bp in length
* 52945 52964: gap of unknown length
* 52965 53539: contig of 575 bp in length
* 53540 53559: gap of unknown length
* 53560 54374: contig of 815 bp in length
* 54375 54394: gap of unknown length
* 54395 55253: contig of 859 bp in length
* 55254 55273: gap of unknown length
* 55274 56527: contig of 1254 bp in length
* 56528 56547: gap of unknown length
* 56548 57376: contig of 829 bp in length
* 57377 57396: gap of unknown length
* 57397 57473: contig of 77 bp in length
* 57474 58337: gap of unknown length
* 58338 58357: gap of unknown length
* 58358 58444: contig of 87 bp in length
* 58445 58464: gap of unknown length
* 58465 59288: contig of 824 bp in length
* 59289 59308: gap of unknown length
* 59309 60499: contig of 1191 bp in length
* 60500 60519: gap of unknown length
* 60520 61859: contig of 1340 bp in length
* 61860 61879: gap of unknown length
* 61880 62752: contig of 873 bp in length
* 62753 62772: gap of unknown length
* 62773 63623: contig of 851 bp in length
* 63624 63643: gap of unknown length
* 63644 64694: contig of 1051 bp in length
* 64695 64715: gap of unknown length
* 64716 65167: contig of 453 bp in length
* 65168 65187: gap of unknown length

```

Query Match      100.0%; Score 393; DB 2; Length 290452;
Best Local Similarity 100.0%; Pred. No. 8.6e-104; Indels 0; Gaps 0;
Matches 393; Conservative 0; Mismatches 0;

QY 1 TTCAACCTCAAACGACGATCGCGATATCAATAAATAAACAAGCAGCAATAGACCGCGTGT 60
Db 53527 TTCAACCTCAAACGACGATCGCGATATCAATAAATAAACAAGCAGCAATAGACCGCGTGT 53468

QY 61 TGAACAACCCGATGCCAGACAAAGTGGTAGATACCTGTTGGCAGACAGTCCCGCGCA 120
Db 53467 TGAACAACCCGATGCCAGACAAAGTGGTAGATACCTGTTGGCAGACAGTCCCGCGCA 53409

QY 121 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCAAAAAATATCGGTGGGAAGCAGCC 180
Db 53407 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCAAAAAATATCGGTGGGAAGCAGCC 53348

QY 181 CAAACACACGATACCGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGCAGG 240
Db 53347 CAAACACACGATACCGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGCAGG 53288

QY 241 CGTTATTATTTTCAGCATCTGTCGCCGCGAGAGAGGCAATGGAAGCCCGCGGAGCAA 300
Db 53287 CGTTATTATTTTCAGCATCTGTCGCCGCGAGAGAGGCAATGGAAGCCCGCGGAGCAA 53228

QY 301 CATTCGCTAGATTTGATTTTAAATATATATATAGCGTAACCTGTTATGCTGTTATATATAT 360
Db 53227 CATTCGCTAGATTTGATTTTAAATATATATATAGCGTAACCTGTTATGCTGTTATATATAT 53168

QY 361 GTGATCTAATACCTTTTAAGTCAGAGTGAATA 393
Db 53167 GTGATCTAATACCTTTTAAGTCAGAGTGAATA 53135

RESULT 4
LOCUS AE016989 292497 bp DNA linear BCT 22-APR-2003
DEFINITION Shigella flexneri 2a str. 2457T section 12 of 16 of the complete genome.
ACCESSION AE016989 AE014073
VERSION AE016989.1 GI:30042681
KEYWORDS
SOURCE Shigella flexneri 2a str. 2457T
ORGANISM Shigella flexneri 2a str. 2457T
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Shigella.
REFERENCE 1 (bases 1 to 292497)
AUTHORS Wei J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W.,
Pournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A.,
Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S.,
Schwartz, D.C. and Blattner, F.R.
TITLE Complete Genome Sequence and Comparative Genomics of Shigella
flexneri Serotype 2a Strain 2457T
JOURNAL Infect. Immun. 71 (5), 2775-2786 (2003)
PUBMED 12704152
REFERENCE 2 (bases 1 to 292497)
AUTHORS Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W.,
Pournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A.,
Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S.,
Schwartz, D.C. and Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
    Location/Qualifiers
    1..292497
        /organism="Shigella flexneri 2a str. 2457T"
        /mol_type="genomic DNA"
        /strain="2457T"
        /serotype="2a"
        /db_xref="taxon:198215"
    98..1169
        /gene="yglJ"
        /locus_tag="S3326"
        /notes="frameshift"

gene

```

```

/pseudo
1186..3537
/gene="yglJ"
/locus_tag="S3327"
1186..3537
/gene="yglJ"
/locus_tag="S3327"
/notes="putative enzyme; Not classified"
/notes="residues 1 to 783 of 783 are 95.65 pct identical to
residues 1 to 783 of 783 from Escherichia coli K-12 :
B3080"
/codon_start=1
/transl_table=11
/product="putative isomerase"
/protein_id="AAP18405.1"
/db_xref="GI:30042682"
/translacion="MKIKTILTPVTCALLISPSAHAANADNYKNVIRNTPGAPQYMKDY
DYDDHQRNPFPLDGNWGHLLPDGNTMGPGVALLTEEVNFWNSNFDLTVWDG
GKYNFTLAYSIPGALVOKLTAKOVQVEMTLERATPRTSLLETKINSKPLDLSGEY
ELLEKEAGEKPLSDKTIAGEPYDQRIKISATRGDKVTFGKVEATWDLTSGSEY
QVHKSPLVQTEINGNRFTSKAHINGSTLTITYSLLHTAQEVSKEQMDRLILAPAF
YLTASQORWEYLKKGLTNPDPQTRVAVKAIETLNGWSPGAVNVTITVSTV
GWFSQNTWPMWDTWKQAFMAHFPIAKENIRAVFSWQIQGDSVFPQDFPDL
IAANLSPERGCGNNBENTKPSLAASWMEVNTVQKWLAEWPKLVAYHDWL
RURDHNGVPYEGATRDKAHNTESGELFTVKGDKEBTQSLNNYARVVEKQVLLG
LEIPQVAASWESGRDDAAVFGIDKEQLDKVANGKSDATVFAENRSODGTLTG
YSLQESVDQASMYSDNHYLAEMATILGKPEAKRYROLAQQLADYINTCMDDPFTQ
FYDVRIEDKPLANGCAGKPIVERGKGPESPLFNGAATQANQADPVVWMLDPKEFN
TFLPLGTAALTNPAGADIYWRGVWDQNFELKGMERYGYRDALADLADTFFRHHK
GLTADGPIONTNPGLTGAQCGAPNFSWSAAHLYMLYNDPFRQ"
3550..3632
/notes="REP (repetitive extragenic palindromic) element,
contains 2 REP sequences"
/gene="yglJ"
/locus_tag="S3328"
3763..5781
/gene="yglJ"
/locus_tag="S3328"
/notes="residues 1 to 672 of 672 are 95.23 pct identical to
residues 1 to 672 of 672 from Escherichia coli K-12 :
B3081"
/codon_start=1
/transl_table=11
/product="putative NADPH dehydrogenase"
/protein_id="AAP18406.1"
/db_xref="GI:30042683"
/translacion="MSYPSLFPAPLDLGFITLKNRVLMGSHHTGLEEYPDGAERLAIFY
AESARHVALIVSGIAPDLTGVMEGAVLNDASQIPHRTITEAVHOGGKIALQI
LHTGRYSYQPHLVAPALQAPINRFPVHELTHEILQILDDFARCAQAREAGYDQVE
VMSEGYLINEFLTNQSRDQMGDYRNRMFAVEVTVTRVGVNDPIIIRLMS
LDLVEGGTFAETVELAEAGATINTGIGWEHARTIATIPVPGAFSWITKRL
KGVSLPLVTNRINDPOVADLISRGDADIVSMARPFADAEALLSKAGSGRADEINT
CIGNCGCLDQIFVGVKTSCLNPRACHETKPEILPAVOKNLAVVAGPAGLAFAIN
AARGHQTFLDHAHEIGGFQNAKQIPGRSEFTETRYRIRMEVITGTLKLNHVT
AAQLQAFDETLASGIVRPVPIGDIDHPKVSYLDLRDKAPVKNVALIGCGGIF
DTMYLSQPGESTSQNIARFCNEWGIDSSIQAGLSQMGQIPRSPQIVMLORQAS
KPGQGLGKTTGTHRTLLSRGVKMPGVSYQKIDDDGLHVGINGTQVLAVDNNVIC
AGQEPNRAQLIDSGKTVHLIGGCDVAMELDARRAIAQGTRLALEI"
complement(5826..6242)
/gene="yglJ"
/locus_tag="S3329"
complement(5826..6242)
/gene="yglJ"
/locus_tag="S3329"
/notes="residues 1 to 138 of 138 are 90.57 pct identical to
residues 1 to 138 of 138 from Escherichia coli K-12 :
B3082"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18407.1"
/db_xref="GI:30042684"

```


AUTHORS	Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M., Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and Hayashi,T.
TITLE	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655
JOURNAL	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
MEDLINE	20557356
PUBMED	11108008
REFERENCE	3
AUTHORS	Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S., Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T., Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasaki,C. and Shinagawa,H.
TITLE	Complete nucleotide sequence of the prophage VTx-Sakai carrying the Shiga toxin I genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak
JOURNAL	Gene 258 (1-2), 127-139 (2000)
MEDLINE	20564182
PUBMED	11111050
REFERENCE	4
AUTHORS	Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K., Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T., Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C., Ogawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and Shinagawa,H.
TITLE	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12
JOURNAL	DNA Res. 8 (1), 11-22 (2001)
MEDLINE	21156231
PUBMED	11258796
REFERENCE	5
AUTHORS	Hattori,M., Ishii,K. and Shiba,T.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUN-2000) Masahira Hattori, Kitasato Institute for Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@genome.lis.kitasato-u.ac.jp) URL:http://genome.lis.kitasato-u.ac.jp/, Tel:81-42-778-8194, Fax:81-42-778-8193)
COMMENT	genome project This work was done in collaboration with Tetsuya Hayashi, Makoto Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Koza Makino, Ken Kurokawa, Katsushi Yokoyama, Masashi Tanaka, Takeshi Honda, Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata (Shinshu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe, Chihiro Sasakawa (University of Tokyo), Hideto Takami (Japan Marine Chihiro Sasakawa (University of Tokyo), Naotake Ogawara (Nara Institute of Science and Technology), Satoru Kuhara (Kyushu University), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science. Location/Qualifiers 1..307962 /organism="Escherichia coli O157:H7" /mol_type="genomic DNA" /strain="O157:H7" /sub_strain="RIMD 0509952" /db_xref="taxon:83334" 72..653 /gene="ECs3910" 72..653 /gene="ECs3910" /note="similar to MDAB_ECOLI gi 1789406 percent identity 100 in 193 aa (Conserved in E.coli K-12)" /codon_start=1 /evidence-not_experimental /transl_table=11 /product="modulator of drug activity B" /protein_id="BAB37333.1" /db_xref="GI:13363383" /translation="MSNTLLINGAKKFAHNSQLNDLTLEVADGTLDLGHIVRA DSDYDKARVONFLVDVVIQMGPMWNGAPWTKYIDDDVTGHGHTLYASDGRTEK MEGSFVPASVGRALFAPENYVQDIRPVLINVLICGLGMVFGAGMAAQSI"
FEATURES	source
gene	gene
CDS	CDS

```
ARNGLADPGLIGVKGCSVALVLIFFQPMGLGMFRPVAGLAGGLVALIVIFCAREI
SRPFLVIGIGVSWFPAAGIGVMTTADVDVDTALMWLSGSHHANMLVIGISAWM
LPAALLLFTPADIALGHQVATGLGVNSRLALLRVAAPILITAVCVSGVNGIF
VGLIAHISRFRILRGQITLLIGVATSGVALLVILADISIGRLAFLPLQUPAGIISLG
GPFLLLLWQRNSF"
complement(3931..4914)
/gene="ECs3915"
complement(3931..4914)
/gene="ECs3915"
/note="probable ferrichrome ABC transporter
(permease), similar to ferrichrome ABC transporters
(permease) e.g. [Synecocystis sp.] gi|7442493|pir||S74438
percent identity 43 in 315 aa, [Bacillus subtilis]
gi|1706795|sp|P49936|PHUB_BACSU percent identity 39"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative permease of ferrichrome ABC
transporter"
/protein_id="BAB37338.1"
/db_xref="GI:13363388"
/translation="WMRTGLVTVIFLALLVGCVVVPGIGARPTAPOTVLOAFLHPEP
QNFENHVIIVLRPLRLAALLGSLGAGALLQAVIRNPLGEPHILGLNAGAAVAV
RASALGLAPVGRPLASTGGALLFILLILSSAGKSGLTVMKVTLCGVVALSAFVSSI
TXAILLDQTLAMRTMLAGLADGWDATIGSAFSLGGVVTALYLAIPASLNMLALG
DRMAQGLSVLRTFTTLLATALLCGAAVSTAGTGFVGLVPOIVRLVWSADLRLV
LPSACVALLLADIARTLTTPHELATGVTALVAGPVFVIMATRMFK"
complement(4911..5720)
/gene="ECs3916"
complement(4911..5720)
/gene="ECs3916"
/note="probable ABC-type iron-siderophore transport system
ATP-binding protein, similar to ABC-type iron-siderophore
Query Match 99.2%; Score 389.8; DB 1; Length 307962;
Best Local Similarity 99.5%; Pred. No. 7.4e-103;
Matches 391; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTCACCTCAACGACAGTCGCGATATCAATAAATAAAGCAATAGAGCGCGGTGT 60
Db 213062 TTCACCTCAACGACAGTCGCGATATCAATAAATAAAGCAATAGAGCGCGGTGT 213121
QY 61 TGAACAAACCGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCAGAGTACCGGCGCA 120
Db 213122 TGAACAAACCGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCAGAGTACCGGCGCA 213181
QY 121 CCAGCCAGAAATCGCCAGTGTATAAAGCAATTCAAAAAATATCGTGGGAGGACAGCC 180
Db 213182 CCAGCCAGAAATCGCCAGTGTATAAAGCAATTCAAAAAATATCGTGGGAGGACAGCC 213241
QY 181 CAACACCAAGTAAACCGGAAACAGACATCTGTGACCTTGGTAAAGAGAGAGCAGG 240
Db 213242 CAACACCAAGTAAACCGGAAACAGACATCTGTGACCTTGGTAAAGAGAGAGCAGG 213301
QY 241 CGTTATTATTTTCAGCATCTGCGCGCAGAGAGGCGATGAAAGCGCGGCGAGACAA 300
Db 213302 CGTTATTATTTTCAGCATCTGCGCGCAGAGAGGCGATGAAAGCGCGGCGAGACAA 213361
QY 301 CATTGCTGTAGATTGATATTAATATATAGCGTAACCTGTATCTGTATCTATTAT 360
Db 213362 CATTGCTGTAGATTGATATTAATATATAGCGTAACCTGTATCTGTATCTATTAT 213421
QY 361 GTGATCAATCACTTTTAAGTCAGAGTGAATA 393
Db 213422 GTGATCAATCACTTTTAAGTCAGAGTGAATA 213454
```

```
RESULT 6
AE016767
LOCUS
DEFINITION Escherichia coli CF073 section 13 of 18 of the complete genome.
ACCESSION AE016767 AE014075
VERSION AE016767.1 GI:26110054
```

KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORSTITLE
JOURNALPUBMED
REFERENCE
AUTHORSTITLE
JOURNALFEATURES
source

gene

CDS

gene

CDS

gene

CDS

```
Escherichia coli CF073
Escherichia coli CF073
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 300359)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence
of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)
12471157
2 (bases 1 to 300359)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 300359
/organism="Escherichia coli CF073"
/mol_type="genomic DNA"
/strain="CF073"
/db_xref="taxon:199310"
188..547
/gene="yqiC"
/locus_tag="c3796"
188..547
/gene="yqiC"
/locus_tag="c3796"
/notes="Escherichia coli K-12 ortholog: b3042; Escherichia
coli O157:H7 ortholog: z4400"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiC"
/db_xref="GI:26110055"
/transl_table=11
/translation="WTYPRTTVALNTYQVRETTMDPKKIEQIAQVHESMPKIGRE
FGEDVBKKIQTQLAQTLRLDLVRSFEEDVQTVLLRTREKLALLQRISELEARYNP
complement(582..788)
/gene="glsS"
/locus_tag="c3797"
complement(582..788)
/gene="glsS"
/locus_tag="c3797"
/function="putative enzyme; Macromolecule synthesis,
modification: Polysaccharides - (cytoplasmic)"
/notes="Escherichia coli K-12 ortholog: b3049; Escherichia
coli O157:H7 ortholog: z4401"
/codon_start=1
/transl_table=11
/product="Glycogen synthesis protein glsS"
/db_xref="GI:26110056"
/transl_table=11
/translation="MLMDHSLNANFDLARSFARMHAEGRPVILAVTGNWDEHR
TWFCARYAYCQMMQTRLELEH"
1052..1681
/gene="yqiJ"
/locus_tag="c3798"
1052..1681
/gene="yqiJ"
/locus_tag="c3798"
/function="putative"
/notes="Escherichia coli K-12 ortholog: b3050; Escherichia
coli O157:H7 ortholog: z4402"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiJ"
```


Db 182461 CGTTATTATTTTCAGCATCTGTCCGCGAGAGGCGATGGAAGCGCGGAGAGCAA 182520

Qy 301 CATTCGCTAGATTGATTTATTAATATTAATAGCTACTGTTATGCTGTTATCTATATTAT 360
|||||

Db 182521 CATTCGCTAGATTGATTTATTAATATTAATAGCTACTGTTATGCTGTTATCTATATTAT 182580
|||||

Qy 361 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 393
|||||

Db 182581 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 182613
|||||

RESULT 7
AE005551
LOCUS
DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 170 of 290.
ACCESSION AE005551
VERSION AE005551.1
KEYWORDS GI:12517850
SOURCE
ORGANISM Escherichia coli O157:H7 EDL933
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 11904)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.
TITLE Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
JOURNAL Nature 409 (6819), 529-533 (2001)
MEDLINE 21074935
PUBMED 11206551
REFERENCE 2 (bases 1 to 11904)
Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
source
location/Qualifiers
1. .11904
/organism="Escherichia coli O157:H7 EDL933"
/mol_type="Genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:155864"
/note="enterohaemorrhagic"
complement(228..620)
/gene="rpsI"
/note="synonym: Z4588"
complement(228..620)
/gene="rpsI"
/note="rpsI"
/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"
/note="Residues 1 to 130 of 130 are 100.00 pct identical to residues 1 to 130 of 130 from Escherichia coli K-12 Strain MG1655: B3230"
/codon_start=1
/transl_table=11
/product="30S ribosomal subunit protein S9"
/protein_id="AAG58358.1"
/db_xref="GI:12517851"
/translation="MAENQYVGTGRKSSAARVFKPQNGKIVNRSLOYPFGRETA RMVVRQPLEVDVMEKLDLYITVKGGISGQAGAIRGHITPMLEYDESURSELKAG FVTRDARQVEKKYGLRKARPPQFSKR"
complement(636..1064)
/gene="rplM"

CDS
/note="synonym: Z4589"
complement(636..1064)
/gene="rplM"
/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"
/note="Residues 1 to 142 of 142 are 100.00 pct identical to residues 1 to 142 of 142 from Escherichia coli K-12 Strain MG1655: B3231"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L13"
/protein_id="AAG58359.1"
/db_xref="GI:12517852"
/translation="MKTFTAKPETVKDYYVDATGKTGLGRLATELARRLRGKHAETPHVDYDGIIVLNADKVAATGNKTDKYVHTHTGGIKQATFEEMIAARRPERVIEIAVKMLPKGLGRAMFRKLKVVAGNEHNAAOQPQLDI"
complement(1283..2410)
/gene="yhcM"
/note="synonym: Z4591"
complement(1283..2410)
/gene="yhcM"
/function="orf; Unknown function"
/note="Residues 1 to 375 of 375 are 100.00 pct identical to residues 1 to 375 of 375 from Escherichia coli K-12 Strain MG1655: B3232"
/codon_start=1
/transl_table=11
/protein_id="AAG58360.1"
/db_xref="GI:12517853"
/translation="MQSVTPTQVLYKALNEGSHQPDVQKEAVSRLEIIYOELINSTPPAPRTSLMARVGLWKGKRDTKHTPVRLYMGVGRGKTWMDLFPQSLPGRKORLHFRPMLRVHERLTALOGOTDPLEIIADRFKAETDVLCEDFEVDITDAMLLGLMLKFAFGITLVATSNIPPDLYNGLQARFLPAIDAIKQHCVMNVADGVDIARLTLTQHLWLSPLHDETRAGMDKLALAGKRENSPTLEINHRPLATGVENQTLAVSTLCVDKASQHDYIALSLFETVLMFDVPMVRLMESEARFIALVDFEYERHVKLVS AEVPLEIYIQCDRLKFEQRCLSRLQEMOSEYLKREHLAG"
2598..3002
/gene="yhcB"
/note="synonym: Z4592"
2598..3002
/gene="yhcB"
/function="orf; Unknown function"
/note="Residues 1 to 134 of 134 are 100.00 pct identical to residues 1 to 134 of 134 from Escherichia coli K-12 Strain MG1655: B3233"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58361.1"
/db_xref="GI:12517854"
/translation="MEWTWEYALIGLVGIIIGAVAMRFGNRKLRQOQALQYELKKNKAEUDYRELVSHFARSABELDTMAHDYQLQCHMAKSSSSLLPELSAEANPFRNLA ESEASNDQAPVQMPRDYSEGSGLLTGAKRD"
3156..4523
/gene="degQ"
/note="synonym: Z4593"
3156..4523
/gene="degQ"
/function="enzyme; Macromolecule degradation: Degradation of proteins, peptides, glyco"
/note="Residues 1 to 455 of 455 are 99.78 pct identical to residues 1 to 455 of 455 from Escherichia coli K-12 Strain MG1655: B3234"
/codon_start=1
/transl_table=11
/product="serine endoprotease"
/protein_id="AAG58362.1"
/db_xref="GI:12517855"
/translation="MKKQTLLSALALSVLTLTSFQAVASIPQGVADQAPLPISLAPMLEKLPVAVSVRVEGTASQKQIPEEFKFFGDDLPDQAPQFPKGLGSGVIINASKGYVLTNNHVINQAKISIQLDNGREFDALKIGSDQSDIALQIQNPSKLTQITADSD KLRVGDFAVAGNPFGLGTATSGIVSALGRSLNLEGLFIQTDAINRNGSGGAL

gene
/note="synonym: Z4599"
complement(636..1064)
/gene="rplM"
/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"
/note="Residues 1 to 142 of 142 are 100.00 pct identical to residues 1 to 142 of 142 from Escherichia coli K-12 Strain MG1655: B3231"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L13"
/protein_id="AAG58359.1"
/db_xref="GI:12517852"
/translation="MKTFTAKPETVKDYYVDATGKTGLGRLATELARRLRGKHAETPHVDYDGIIVLNADKVAATGNKTDKYVHTHTGGIKQATFEEMIAARRPERVIEIAVKMLPKGLGRAMFRKLKVVAGNEHNAAOQPQLDI"
complement(1283..2410)
/gene="yhcM"
/note="synonym: Z4591"
complement(1283..2410)
/gene="yhcM"
/function="orf; Unknown function"
/note="Residues 1 to 375 of 375 are 100.00 pct identical to residues 1 to 375 of 375 from Escherichia coli K-12 Strain MG1655: B3232"
/codon_start=1
/transl_table=11
/protein_id="AAG58360.1"
/db_xref="GI:12517853"
/translation="MQSVTPTQVLYKALNEGSHQPDVQKEAVSRLEIIYOELINSTPPAPRTSLMARVGLWKGKRDTKHTPVRLYMGVGRGKTWMDLFPQSLPGRKORLHFRPMLRVHERLTALOGOTDPLEIIADRFKAETDVLCEDFEVDITDAMLLGLMLKFAFGITLVATSNIPPDLYNGLQARFLPAIDAIKQHCVMNVADGVDIARLTLTQHLWLSPLHDETRAGMDKLALAGKRENSPTLEINHRPLATGVENQTLAVSTLCVDKASQHDYIALSLFETVLMFDVPMVRLMESEARFIALVDFEYERHVKLVS AEVPLEIYIQCDRLKFEQRCLSRLQEMOSEYLKREHLAG"
2598..3002
/gene="yhcB"
/note="synonym: Z4592"
2598..3002
/gene="yhcB"
/function="orf; Unknown function"
/note="Residues 1 to 134 of 134 are 100.00 pct identical to residues 1 to 134 of 134 from Escherichia coli K-12 Strain MG1655: B3233"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58361.1"
/db_xref="GI:12517854"
/translation="MEWTWEYALIGLVGIIIGAVAMRFGNRKLRQOQALQYELKKNKAEUDYRELVSHFARSABELDTMAHDYQLQCHMAKSSSSLLPELSAEANPFRNLA ESEASNDQAPVQMPRDYSEGSGLLTGAKRD"
3156..4523
/gene="degQ"
/note="synonym: Z4593"
3156..4523
/gene="degQ"
/function="enzyme; Macromolecule degradation: Degradation of proteins, peptides, glyco"
/note="Residues 1 to 455 of 455 are 99.78 pct identical to residues 1 to 455 of 455 from Escherichia coli K-12 Strain MG1655: B3234"
/codon_start=1
/transl_table=11
/product="serine endoprotease"
/protein_id="AAG58362.1"
/db_xref="GI:12517855"
/translation="MKKQTLLSALALSVLTLTSFQAVASIPQGVADQAPLPISLAPMLEKLPVAVSVRVEGTASQKQIPEEFKFFGDDLPDQAPQFPKGLGSGVIINASKGYVLTNNHVINQAKISIQLDNGREFDALKIGSDQSDIALQIQNPSKLTQITADSD KLRVGDFAVAGNPFGLGTATSGIVSALGRSLNLEGLFIQTDAINRNGSGGAL

gene
/note="synonym: Z4599"
complement(636..1064)
/gene="rplM"
/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"
/note="Residues 1 to 142 of 142 are 100.00 pct identical to residues 1 to 142 of 142 from Escherichia coli K-12 Strain MG1655: B3231"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L13"
/protein_id="AAG58359.1"
/db_xref="GI:12517852"
/translation="MKTFTAKPETVKDYYVDATGKTGLGRLATELARRLRGKHAETPHVDYDGIIVLNADKVAATGNKTDKYVHTHTGGIKQATFEEMIAARRPERVIEIAVKMLPKGLGRAMFRKLKVVAGNEHNAAOQPQLDI"
complement(1283..2410)
/gene="yhcM"
/note="synonym: Z4591"
complement(1283..2410)
/gene="yhcM"
/function="orf; Unknown function"
/note="Residues 1 to 375 of 375 are 100.00 pct identical to residues 1 to 375 of 375 from Escherichia coli K-12 Strain MG1655: B3232"
/codon_start=1
/transl_table=11
/protein_id="AAG58360.1"
/db_xref="GI:12517853"
/translation="MQSVTPTQVLYKALNEGSHQPDVQKEAVSRLEIIYOELINSTPPAPRTSLMARVGLWKGKRDTKHTPVRLYMGVGRGKTWMDLFPQSLPGRKORLHFRPMLRVHERLTALOGOTDPLEIIADRFKAETDVLCEDFEVDITDAMLLGLMLKFAFGITLVATSNIPPDLYNGLQARFLPAIDAIKQHCVMNVADGVDIARLTLTQHLWLSPLHDETRAGMDKLALAGKRENSPTLEINHRPLATGVENQTLAVSTLCVDKASQHDYIALSLFETVLMFDVPMVRLMESEARFIALVDFEYERHVKLVS AEVPLEIYIQCDRLKFEQRCLSRLQEMOSEYLKREHLAG"
2598..3002
/gene="yhcB"
/note="synonym: Z4592"
2598..3002
/gene="yhcB"
/function="orf; Unknown function"
/note="Residues 1 to 134 of 134 are 100.00 pct identical to residues 1 to 134 of 134 from Escherichia coli K-12 Strain MG1655: B3233"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58361.1"
/db_xref="GI:12517854"
/translation="MEWTWEYALIGLVGIIIGAVAMRFGNRKLRQOQALQYELKKNKAEUDYRELVSHFARSABELDTMAHDYQLQCHMAKSSSSLLPELSAEANPFRNLA ESEASNDQAPVQMPRDYSEGSGLLTGAKRD"
3156..4523
/gene="degQ"
/note="synonym: Z4593"
3156..4523
/gene="degQ"
/function="enzyme; Macromolecule degradation: Degradation of proteins, peptides, glyco"
/note="Residues 1 to 455 of 455 are 99.78 pct identical to residues 1 to 455 of 455 from Escherichia coli K-12 Strain MG1655: B3234"
/codon_start=1
/transl_table=11
/product="serine endoprotease"
/protein_id="AAG58362.1"
/db_xref="GI:12517855"
/translation="MKKQTLLSALALSVLTLTSFQAVASIPQGVADQAPLPISLAPMLEKLPVAVSVRVEGTASQKQIPEEFKFFGDDLPDQAPQFPKGLGSGVIINASKGYVLTNNHVINQAKISIQLDNGREFDALKIGSDQSDIALQIQNPSKLTQITADSD KLRVGDFAVAGNPFGLGTATSGIVSALGRSLNLEGLFIQTDAINRNGSGGAL

LNMLGELICINTAILAPGGSGVIGFPAIPSNMARTLAQQLIDFGEIKRGLIGKTEM
SADIARFNLDVQRFVSEVLPSSGSAKAGVXTGDIITSLNGKPLNSPABLSRIAT
TEPGTKVGLDQKPLEVEVITLDTSSASAEIMITPALEGATLSDGQKDGKGI
KIDEVKGSPPAQAGLQKDDVIIIVNRDRVNSIAEMRKRVLAAKPAIIALQIVRNE
YLLMR"
4613..5680
/gene="degs"
/note="synonym: Z4594"
4613..5680
/gene="degs"
/function="enzyme; Macromolecule degradation: Degradation
of proteins, peptides, glyco"
/note="Residues 1 to 355 of 355 are 99.71 pct identical to
residues 1 to 355 of 355 from Escherichia coli K-12 Strain
MG1655: B3235"
/codon_start=1
/transl_table=11
/product="protease"
/protein_id="AAG58363.1"
/db_xref="GI:12517856"
/translation="MFVKLRSVAITGLIVGAILLVAMPRLSLNPLSTPQFDSTDETP
ASYNLAARPAVAVVNYRGLNTSHNQLRIETLGSVINDQRYIITNKHVINDAD
QIIVAQDQRFVALLVGSLSLTDJAVLKINATGGLPTIIPINARVPHIGDVLAIGN
PYNKQITIGIISATIGLNTQRPLOQFNLDASINHGSGALVSLGELMGINIL
SPDKNDGTPSGIIFAIPTFQATIKMDKIRGVIRGYIGIGREIAPLHAQGGGI
DOLQGIIVNEVSPDGPANAGIQVNDLIISVDNKPASIALETMDQVAEIRPGSVIPV
VMRDKQLTQVIOEYPATN"
complement(5744..5682)
/gene="mdh"
/note="synonym: Z4595"
complement(5744..5682)
/genes="mdh"
/function="enzyme; Energy metabolism, carbon: TCA cycle"
/note="Residues 1 to 312 of 312 are 100.00 pct identical
to residues 1 to 312 of 312 from Escherichia coli K-12
Strain MG1655: B3236"
/codon_start=1
/transl_table=11
/product="malate dehydrogenase"
/protein_id="AAG58364.1"
/db_xref="GI:12517857"
/translation="MKVAVLGAAGIGQALALLKLTOLPSGSELSLYDIAPVTPGVAV
DUSHPTAKVIGFSGEDATPALEGADVLSAGVARKGMDRSLFNVAGLVKNLV
QQVAKTCPKACIGIINTVNTVAIAEVLKAGVYDKNKGFTVILDIIRSNFVAE
LKQKQGEVEVPVIGSGVTIPLLSQVPGFSFTEQEVADLTKRIAGTEVBEAKA
GGGATLSMGAAARFGLSLVALQEOGVVEGQVAFPSQFLLGKNGVE
ERKSICITLSAFQNALLEGMLDLTKKDIALGEEFVNK"
7117..7587
/gene="argR"
/note="synonym: Z4596"
7117..7587
/genes="argR"
/function="regulator; Amino acid biosynthesis: Arginine"
/note="Residues 1 to 156 of 156 are 100.00 pct identical
to residues 1 to 156 of 156 from Escherichia coli K-12
Strain MG1655: B3237"
/codon_start=1
/transl_table=11
/product="repressor of arg regulon; cer-mediated site
specific recombination"
/protein_id="AAG58365.1"
/db_xref="GI:12517858"
/translation="MRSSAKQELVKAFLKKEKFSOGHEIVAAQOQGFNDINQS
KVRMLTKFVAVIRAKMEMVYCLPAELGVPFTSSPLNVLVDIDYNDVAVVHTSP
GAQQLARLDSLGKAGILGTIAGDITFTTPANGFTVKDLYEAILLEDFDEL"
7901..8215
/gene="yhcN"
92.3%; Score 362.8; DB 1; Length 11904;
Best Local Similarity 95.5%; Pred. No. 4.8e-95;
Matches 364; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 TTCAACCTCAACGACAGTCGCGATATCAATAAAACAGACGCAATAGAGCGCGTGT 60

Db 11539 TTCAACCTCAACGACAGTCGCGATATCAATAAAACAGACGCAATAGAGCGCGTGT 11598
Qy 61 TGAACACCGCGGATGCCAGACAAAGTCTAGATACCTCTTGGCACAAGTACCCGCGCA 120
Db 11599 TGAACACCGCGGATGCCAGACAAAGTCTAGATACCTCTTGGCACAAGTACCCGCGCA 11658
Qy 121 CCAGCCAGAAATTCGCCAGTGTATAAAAGCAATTCAAAAATATCGGTGGAAAGGACAGCC 180
Db 11659 CCAGCCAGAAATTCGCCAGTGTATAAAAGCAATTCAAAAATATCGGTGGAAAGGACAGCC 11718
Qy 181 CAAACACCGAGTAAACGGGAACAGACTCATGTGTACCTTGGTTGTAAAGAGAGAGCAGG 240
Db 11719 CAAACACCGAGTAAACGGGAACAGACTCATGTGTACCTTGGTTGTAAAGAGAGAGCAGG 11778
Qy 241 CGTTATTATTTTCAGCATCTGTCCCGCAGAGAGGCGCATGGAAGCCCGGCGAGAGCAA 300
Db 11779 CGTTATTATTTTCAGCATCTGTCCCGCAGAGAGGCGCATGGAAGCCCGGCGAGAGCAA 11838
Qy 301 CATTCGTAGATGATATTAATATATTAGCTACTGTATCTGTTATCTATATTAT 360
Db 11839 CATTCGTAGATGATATTAATATATTAGCTACTGTATCTGTTATCTATATTAT 11898
Qy 361 GTGATC 366
Db 11899 GTGATC 11904
RESULT 8
AE008855 21252 bp DNA linear BCT 23-APR-2003
LOCUS Salmonella typhimurium LT2, section 159 of 220 of the complete
DEFINITION Genome.
ACCESSION AE008855 AE006468
VERSION AE008855.1 GI:16421920
KEYWORDS
SOURCE Salmonella typhimurium LT2
ORGANISM Salmonella typhimurium LT2
REFERENCE
AUTHORS
McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,
Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F.,
Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,
Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W.,
Stonking, T., Nhan, M., Waterston, R. and Wilson, R.K.
Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2
JOURNAL Nature 413 (6858), 852-856 (2001)
MEDLINE 21534948
PUBMED 11677609
REFERENCE 2 (bases 1 to 21252)
AUTHORS
The Salmonella typhimurium Genome Sequencing Project
DIRECT SUBMISSION
Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT Supported by NIH grant 5U 01 AI43283
Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and
GeneMark; <http://opal.biology.gatech.edu/GeneMark/>
EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>,
and Pedro Romero and Peter Karp at EcoCyc;
<http://ecocyc.PangeaSystems.com/ecocyc/>
The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Vides and
ReguonDB;

http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

```
FEATURES
    source
        1..21252
            /organism="Salmonella typhimurium LT2"
            /mol_type="genomic DNA"
            /strain="LT2; SGSC 1412; ATCC 700720"
            /db_xref="ATCC:700720"
            /db_xref="taxon:99287"
            /note="LT2"
    Gene
        207..482
            /gene="yhcn"
            /note="synonym: STM3361"
    RBS
        207..212
            /gene="yhcn"
            /note="putative RBS for yhcN; RegulonDB:STMSIH003374"
    CDS
        219..482
            /gene="yhcn"
            /note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 73% identity in aa 18 - 104"
            /codon_start=1
            /transl_table=11
            /product="putative outer membrane protein"
            /protein_id="AAL22230.1"
            /db_xref="GI:16421921"
            /translation="MKIKTTVATLSILSVLSFGFAEPISAEQAQNREATIGSVSVA
            IGSSPMDNMAMLSKXADEQATAYHTEARSGSNWHATAELYK"
    Gene
        579..852
            /gene="STM3362"
    RBS
        579..584
            /gene="STM3362"
    CDS
        586..852
            /note="putative RBS for STM3362; RegulonDB:STMSIH003375"
            /gene="STM3362"
            /note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 53% identity in aa 18 - 104"
            /codon_start=1
            /transl_table=11
            /product="putative periplasmic protein"
            /protein_id="AAL22231.1"
            /db_xref="GI:16421922"
            /translation="MKTKYIIASLGATLLSFGANAAVHGVNAEQANLQPMGTISVS
            QIGSTPMROEIVAKAKAGANSYRIIELKEDGNWHATAELYK"
    gene
        complement(912..1197)
            /gene="yhco"
            /note="synonym: STM3363"
            /complement(912..1184)
    CDS
        /gene="yhco"
            /note="similar to E. coli orf, hypothetical protein (AAC76271.1); Blastp hit to AAC76271.1 (90 aa), 76% identity in aa 1 - 90"
            /codon_start=1
            /transl_table=11
            /product="putative cytoplasmic protein"
            /protein_id="AAL22232.1"
            /db_xref="GI:16421923"
            /translation="MNVYTFDFNDIKQSDFYREFTOTFGLASBKVSDLDLTLDVMS
            DILPLLEIFVHLDPKURRYGALILLFDEAELEGRLFRNVRH"
    RBS
        complement(1192..1197)
            /gene="yhco"
            /note="putative RBS for yhcO; RegulonDB:STMSIH003376"
            /complement(1356..1323)
    gene
        /gene="yhcp"
            /note="synonym: STM3364"
            /complement(1356..1323)
    CDS
```

```

            /gene="yhcp"
            /note="similar to E. coli orf, hypothetical protein (AAC76272.1); Blastp hit to AAC76272.1 (655 aa), 92% identity in aa 1 - 555"
            /codon_start=1
            /transl_table=11
            /product="putative inner membrane protein"
            /protein_id="AAL22233.1"
            /db_xref="GI:16421924"
            /translation="MGIFSIANOHIRFAVKACAIVLALFIFGHFOLETPRMAVLTA
            IVAAGPAPAGGPPSGAIRYRGMRLIIGTFICIAALIIISIMRAPLLMILVCCVW
            AGFTWISSIVRIENSYAGLSGYTALIIVITQTPELTPQALPQALPQALPQALPQ
            ILADLLPSPSRIKQVDRELDLSVAQYQMLQCIKRGSEEDVNDAGDLVRVRAALE
            GMRNLNMESSWRVRANRKLALNTLSLTITQSCETYLITQNTPELITDTFRELFET
            FETIQDVHQRKMRVIVTGERETPTVLSWVGATRYLLKKGVSISNTKISATE
            EEILQGPVVVKVESAEERHAMVNFRTLLSCILGTILMTWGTSGNVMIAVVT
            LAMRLPNRMVCIDFIYGLAALPLGLYPLIIPNTQOSMLLCLSLAVLGFPIGIE
            VOKRGLSGALASTINIIVLDNPLMTFHSOFDLSALGOIVGCMIAFIVILVRDKSK
            DRTGRVLLNQFVSAASANTNVVRKRNRLPALYQQLFLLMKNKFGDLPKPRALATM
            ITAQRDRAPIPVEDLSVFRHQRARTADHVISAGSDDKRRRIFGQLLDELDIYQEK
            LRIWEAPPQVTEPKSLTGLMKHYQNALTDS"
            /complement(3329..4270)
            /gene="yhco"
            /note="synonym: STM3365"
            /complement(3329..4261)
            /gene="yhco"
            /note="similar to E. coli putative membrane protein (AAC76273.1); Blastp hit to AAC76273.1 (310 aa), 92% identity in aa 1 - 310"
            /codon_start=1
            /transl_table=11
            /product="putative membrane located multidrug resistance protein"
            /protein_id="AAL22234.1"
            /db_xref="GI:16421925"
            /translation="MKTLRKLSRTAITLVILAFIAIFRAWVYTESPWTRDARFS
            ADVAIAPVAGLITHVNVHDNOLVKKQDVLFTIDQRYKALAEAEADVAVYQVLAQ
            EKQEGARRNLGQVMSREEDQANNVLOTVLHQAKAQATREDLAKLDERTVIRAP
            ADGWNTLVNAYAGEFTTRGSTAVLVKMSFPYQVQWMEETKLEGVPGVRAETPLGS
            NRVLKTVDSVAAAGVNASDSDAKGNATIDSNLEWRLAQRVFRIRLDEQQGNLWP
            AGTATVITGKQDRDASQDSFFRKLHLRLRFBG"
            /complement(4265..4270)
            /gene="yhco"
            /note="putative RBS for yhcQ; RegulonDB:STMSIH003377"
            /complement(4269..4481)
            /gene="yhcr"
            /note="synonym: STM3366"
            /complement(4269..4472)
            /gene="yhcr"
            /note="similar to E. coli orf, hypothetical protein (AAC76274.1); Blastp hit to AAC76274.1 (90 aa), 98% identity in aa 24 - 90"
            /codon_start=1
            /transl_table=11
            /product="putative inner membrane protein"
            /protein_id="AAL22235.1"
            /db_xref="GI:16421926"
            /translation="MSLFPVIVVFGLSFPPIFFELLLSLAFWLVRMLVPTGIYDFV
            WHPALENTALYCCLFYLIISRLSV"
            /complement(4476..4481)
            /gene="yhcr"
            /note="putative RBS for yhcR; RegulonDB:STMSIH003378"
            /complement(4654..5583)
            /gene="yhcs"
            /note="synonym: STM3367"
            /complement(4654..5583)
            /gene="yhcs"
            /note="similar to E. coli putative transcriptional regulator LysR-type (AAC76275.1); Blastp hit to AAC76275.1 (309 aa), 95% identity in aa 1 - 309"
            /codon_start=1
            /transl_table=11
            /product="putative LysR family transcriptional regulator"
```


Orthologue of E. coli ygiC (YGIC_ECOLI); Pasta hit to YGIC_ECOLI (386 aa), 94% identity in 386 aa overlap

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07714.1"
/db_xref="GI:16504266"
/db_xref="TrEMBL:Q8FFP6"
/translation="MERSVITRPPDRKATYGFNFTMYGEPYWCEDAYVYKLTAAQ
VEKLEDTAELQCMCKLVKVERIYASDELMTKFIKPHTWGVRVQSMOTQPSLSRLD
LAWDGEPEKLENNADPTSLYEAFAQWLEQINAGNLPEGSDQPSLOEKLE
RAEUREQVQLHLHLCGRDIVERDRTIYQDCARAEATIEFYIDDIQGLGKQ
FIDQDVANFLYKLPWFEMFEMFSTKLEDAWVLEPAWKSIIISNKALLPLLWEM
FPHNLNLAFAEBEHPMDKYVVKPIFREGANVSIENGKTTIESYEGYGEEMI
VQQFPLPFGDSYTLIGSLINDQAPGIGIREDRALITQLSREYPIHFVEG"
complement (3659..4489)
/gene="STY3367"
complement (3659..4489)
/gene="STY3367"
/note="Orthologue of E. coli ygiD (YGID_ECOLI); Pasta hit
to YGID_ECOLI (271 aa), 87% identity in 266 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07715.1"
/db_xref="GI:16504267"
/db_xref="GOA:Q8Z3N7"
/db_xref="TrEMBL:Q8Z3N7"
/translation="MVMILTCLSCDKNIMSLTCLPALFLHGSPMNVLDNDYTRAWR
RLCEALPRAIIVVSAHWYTRGTGTAMERPTLHDGCGPQALDYDHYAPGPSAL
AORVLELLAPVALDKAWGFDHGSWGLIKWYNADIPMVLSDVSTKPAKHFE
GRKALRDEGVMLVAGSNVNIETRVHGNINIPYPAASFNDRFVKANLWQGPVEQ
HPLVNLQHEGALLNPTEHFLPLLYVLGAWDGKEPITIPVDGIEMSGISMLSVQVG
"
4588..5361
/gene="STY3368"
4588..5361
/gene="STY3368"
/note="Orthologue of E. coli ygiE (YGIE_ECOLI); Pasta hit
to YGIE_ECOLI (257 aa), 93% identity in 257 aa overlap.
Contains multiple possible membrane spanning hydrophobic
domains"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD07716.1"
/db_xref="GI:16504268"
/db_xref="GOA:Q8XGR4"
/db_xref="TrEMBL:Q8XGR4"
/translation="MSVPLILTLAAGATFAGFLGVLGKPSNRVLAFLSFGAAGIM
LLISLMELPALDTEGMSPLGYGMFIIGLLGYGLDELLPHAPQDLVQKQOPLP
GSIKETALLITIGLSHPEPIATFVTASSNLELGFGLALVALHNIPEGLAVAGPV
YAATGSKRTAIFWAGISGNAEILGGVLAWLILGSLVSPVVAANAAGVAGIIVALSVD
ELMPLAKEIDPNPNPSYGLGMSINGLSLVLQITIGIG"
6004..7800
/gene="STY3370"
/note="synonym: asst"
6004..7800
/gene="STY3370"
/note="Similar to Klebsiella pneumoniae arylsulfate
sulfotransferase asst TR:P97036 (EMBL:U32616) (598 aa)
fasta scores: E(): 0, 86.8% id in 598 aa"
/codon_start=1
/transl_table=11
/product="probable arylsulfate sulfotransferase"
/protein_id="CAD07717.1"
/db_xref="GI:16504269"
/db_xref="GOA:Q8Z3N6"
/db_xref="TrEMBL:Q8Z3N6"
/translation="MFDQVRKTIILAGAVALTCLTAASTFAAGFQAPQAGKIGAVVV
DPVGNAPITVYLDVSHIISDVKVTVHGKGVPTVTVGKESLETVDG:PIGLYQ
KFNWVTVYKENGKMKDDVYQTSAINVHMDNRSIDLQQTIVIKVAGFEDRLY
LVNTHTFQGEFPHGKDKNAGILDAGPAGGALPFDIATPITFVVDITQGEYRWLID

gene

CDS

QDTFYNGHDMNINKRGYLMGIRETPRGTTTAVOQHWYEFDMVMGOILADHKLPRGFID
ASHESIETVNGTVLLRVGRKDYRKEDGIHVHTIRDOIIEVDKSRGVVVDVNDLTKILD
MRDALLGALDAGAVCVNVDLAHQQAOKLSPDTPYDALGVAGRWNAHVSIAVDK
DDSIILSRHQGIVKIGRQKQWILAPSKGNKQKASKLLKPKVDHDKGKPLTCDENGK
KDTDPDFTTQHTAMRSSKGLTIVDNGRGEQEPALFETKPIKSVFVETKIDKGGT
VQVWEYKGRGDFYFSPITSVVEYQKDRMTFGGSGINLFDVGKPTVGLKNEIDYK
TKYKVEIDVLSLQPNQTHYRALLVHPQTQMKF"
7820..8491
/gene="dsbA"
/note="synonym: STY3371"
7820..8491
/gene="dsbA"
/note="similar to Enterobacter amnigenus disulfide
isomerase dsbA TR:Q9XDP1 (EMBL:AF012826) (222 aa) fasta
scores: E(): 0, 90.1% id in 222 aa"
/codon_start=1
/transl_table=11
/product="disulfide isomerase"
/protein_id="CAD07718.1"
/db_xref="GI:16504270"
/db_xref="GOA:Q8Z3N5"
/db_xref="TrEMBL:Q8Z3N5"
/translation="MSKWTITSLFKSVLTAALVTPFAASAFETGTDYMLEKPIPA
DKTLIKVSYACPCPKYKDKAVTGVSDKVADIVTPTPEHLETKGKQKASVPAVL
IAKDRAAGISLFDPAKSOFKKAFWYTAHYDKERWSDKDPAAFIKTGLDAGMSA
DFAALKDPAPVQETLEKWKAAAYDVAKIQGVPAIVVNGKYLIIYTKNIKSIDMAELVRE
LATKK"
7847..8488
misc_feature
7847..8488
Query Match 73.7%; Score 289.8; DB 1; Length 258050;
Best Local Similarity 85.2%; Pred. No. 1:5e-73;
Matches 335; Conservative 0; Mismatches 57; Indels 1; Gaps 1;
QY 1 TTCAACTCAAACGACAGTCGCGATATCAAAATAAAACAAGCAGCAATAGACGCGGTGT 60
DB 180504 TTCAACTCAAACGACAGCGCGATATTAAATAAAACAAGCAGCAATAGACGCGGTAT 180563
QY 61 TGAACACGCGGATCCGACAGCAAGTCGTAGTACCTGTGCGACAGTACCGCGCGCA 120
DB 180564 TGAACACGCGGATCCGACAGCAAGTCGTAGTACCTGTGCGACAGTACCGCGCGCA 180623
QY 121 CCAGCCAGAAATCGCCAGTCGATATAAAAGCAATCAAAATAATCGTGGGAAGCAGACCC 180
DB 180624 CCAGCCAGAAATAGCCAGTCGACAGCAATTCGAAATAATCGTGGGAAGCAAGAC 180683
QY 181 CAACACACAGATAACCGGAAACAGACTCATGTGTGACCTTGGTTGTAAGAGAGAGCAGG 240
DB 180684 CAACACACAGATAACCGGAAACAGACTCATGTGTGACCTTGGTTGTAAGAGAGAGCAGG 180743
QY 241 CGTTATTATTTCAGCATCTGTCGCGCAGAGAGGSCATGGAAAGCCGCGCGAGAGCAA 300
DB 180744 CGATAGATTGTTAGC-TACATCGCCGACAGAGAGGSCAGAGACGCCGCGAGAGCGCA 180802
QY 301 CATTCGTGTAGATTGATTAATTAATATATTAGCGTAACCTGTTATGCTGTTATCTATATTAT 360
DB 180803 CATAGCTCTTATTTCGAATAATAGTATTAAACGTAATGTTGTTGCTGTTATCTATATTAT 180862
QY 361 GTGATCTAAATCACTTTTAACTAGTCAGAGTGAATA 393
DB 180863 GTGATCTAAATCACTTTTAACTAGTCAGAGTGAATA 180895
RESULT 10
AE016845
LOCUS
DEFINITION Salmonella enterica subsp. enterica serovar typhi Ty2, section 12
of 16 of the complete genome.
ACCESSION AE016845 AE014613
VERSION AE016845.1 GI:29139182
KEYWORDS
SOURCE Salmonella enterica subsp. enterica serovar Typhi Ty2
ORGANISM Salmonella enterica subsp. enterica serovar Typhi Ty2
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Salmonella.
 1 (bases 1 to 299991)
 Deng W., Liou S.-R., Plunkett III, G., Mayhew, G.F., Rose, D.J.,
 Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R.
 Comparative Genomics of Salmonella enterica Serovar Typhi Strains
 Ty2 and Ctr18
 J. Bacteriol. 185 (7), 2330-2337 (2003)
 22531367
 12644504
 2 (bases 1 to 299991)
 Deng W., Liou S.-R., Plunkett, G. III, Mayhew, G.F., Rose, D.J.,
 Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R.
 Direct Submission
 Submitted (25-SEP-2002) Laboratory of Genetics, University of
 Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
 Location/Qualifiers
 1. .299991
 /organism="Salmonella enterica subsp. enterica serovar
 Typhi Ty2"
 /mol_type="genomic DNA"
 /strain="Ty2"
 /serovar="Typhi"
 /sub_species="enterica"
 /db_xref="taxon:209261"
 complement (83..709)
 /gene="ftsJ"
 /locus_tag="t3214"
 complement (83..709)
 /gene="ftsJ"
 /locus_tag="t3214"
 /notes="corresponds to STV3475 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="cell division protein"
 /protein_id="AAO70750.1"
 /db_xref="GI:29139183"
 /translation="MTGKRSASRRWLQEHFSDKYVOAOKGLRSRAWFKLDEIQ
 SDKLPKGVTVVDLGAAPGQSVYVTOIGKGRILACDLLPMDPIVGVDFLOQDFRD
 ELVKALLREVDGSKVQVYMSDAPNMSGTRAVDIPRAMYLVELALDNCBVLAPGGS
 FVYKFGSGDFEYREIRLSLFTKVKVRKPSRRSRREVIYATGRK"
 797..1129
 /locus_tag="t3215"
 797..1129
 /locus_tag="t3215"
 /notes="corresponds to STV3476 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAO70751.1"
 /db_xref="GI:29139184"
 /translation="WTRFQSQKQKVTWNLSTKOKHLKGLAHP:KPVVMLGNNGLTG
 GYLAHEEQALEHHELIKVLASDEIRETKLIVDAIVRTGACNVQVIGKTLVLRPTK
 ERKISLPR"
 complement (1285..1761)
 /gene="greA"
 /locus_tag="t3216"
 complement (1285..1761)
 /gene="greA"
 /locus_tag="t3216"
 /notes="corresponds to STV3477 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="transcription elongation factor"
 /protein_id="AAO70752.1"
 /db_xref="GI:29139185"
 /translation="MQAIPMTLRGAELKEEEDFLKSVRRPEIIAIAEAREHGDJLKE
 NABYHAARQQGCEGRKIDIEAKLSNAQVIDTKPNNRVRVIGATVVLNLDDEE
 QTVRIYGDDEADFKQNLISVNSPIARGLGKQDDVWIKTPGDDVEYELKVEYL"
 2009..3442
 /locus_tag="t3217"
 2009..3442
 /locus_tag="t3217"

CDS
 2009..3442
 /locus_tag="t3217"
 /notes="corresponds to STV3479 from Accession AL513382:
 Salmonella typhi Ctr18; D-alanyl-D-alanine
 carboxypeptidase"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein"
 /protein_id="AAO70753.1"
 /db_xref="GI:29139186"
 /translation="MRFSRFIIGLTTSIAFSVQAANIDEYIKQIPAGANLALMVQKIG
 APAPDHYHQWALPASTQKVTALAAALIQGPDPFRFTTILETKGVNDGILKGLDI
 ARFGDPTLRRODIRNMVATLKSGVTQIDGNVLDITSFASHDKAPGWNLDLQCF
 SAPAAAIYDRNCFSVLSAQKPNDLARIVASYVPTMFQSVRLTPGSDADAOYCE
 LDVVPGLMRYLTGCLPORDPLPAFAIQDGSYAGALLKQELKEAGITVRGTLIR
 QTVNEFGTIVASKQSAFLHDLKIMLKSDNMIADTVFRMIGHVFNFRVFGWRAGSD
 AVRIIRQAGIDIGNTIADGSGLSRHNLIAPTNMVQLQYIAQHDNENLFIIMPL
 AGYDGLQYRAGLHQAGVCKVSAKTSGLQGVYNLAGFITTAGSGORMAFVQLSGYAV
 PPADQRNRIPLVRFESRLVKDIYQNN"
 complement (3573..4745)
 /locus_tag="t3218"
 complement (3573..4745)
 /locus_tag="t3218"
 /notes="corresponds to STV3480 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="probable GTP-binding protein"
 /protein_id="AAO70754.1"
 /db_xref="GI:29139187"
 /translation="MKFVDEASILVVAGDGGNGCVSRREKYPKGGDGGDGGD
 VVMEADENLTLIDYRFEKSFRAERGONGASRDCTGKRGKVTIKVPGVTRVDOGTG
 ETWGDMLTKHGRLLVAKGWHGLGNTRFKSVNVRTPKNTGTPGDKRDLLEMLLA
 DVGMLGNPAGSTFIRAVSAAPKVADYPTTLVPSLGVRMDSEKSPVADIPGLI
 EGAEAGLGIREFLKHLCRCVRLHLLDIDPDGSDPVENARIIGELKYSQDLAAK
 PRWLNFNKIDLMKDSAEBAKAIABALGWEKGIYILSAASQGVKDLCDWNTFFLE
 NP:AAQEAERKQPEVFEMKADYHQQLAEVSDADDDDDDDDEDEGEVEFIYKR"
 complement (4761..5726)
 /locus_tag="t3219"
 complement (4761..5726)
 /locus_tag="t3219"
 /notes="corresponds to STV3481 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="putative membrane protein"
 /protein_id="AAO70755.1"
 /db_xref="GI:29139188"
 /translation="MQQAGIGILLALTTCMCWGPALPIAMKQVLEWEPSTIVYRFL
 MASIGILAVLRKLPPIRFRKRWLVLAICGLFNFTLSPSSQVLSPTASQ
 VIGQLSPGVWASVFILKEKRGQTVIGALMLSLGVNFTSLIEIFRDLTPWG
 VIFGVGAAMVWSYGVAAQVLLRRLASQOILFLYLTCTIALPLAKPVIQALSDWQ
 LACILFCGNTLVYGCALAEAMARQAAQVSAIITLPLTLLFSLLSWAMPDFPAR
 PMLNLLGLGAFVUVAGAMYSAIGHRIWGLRKHTVVSQPRSGE"
 complement (5856..6113)
 /gene="rpnA"
 /locus_tag="t3220"
 complement (5856..6113)
 /gene="rpnA"
 /locus_tag="t3220"
 /notes="corresponds to STV3482 from Accession AL513382:
 Salmonella typhi Ctr18"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal subunit protein L27"
 /protein_id="AAO70756.1"
 /db_xref="GI:29139189"
 /translation="MAHKKAGSGSTRNGRDSKRLGVRKRGGEAVLAGSIIVRQRTK
 FHAGTVGGRDHTLFAKADGVKPKVKGPKRKYISIAVE"
 complement (6133..6444)
 /gene="rplU"
 /locus_tag="t3221"
 complement (6133..6444)


```

/genes="rplU"
/locus_tag="t3221"
/notes="corresponds to STY3483 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L21"
/protein_id="AA070757.1"
/db_xref="GI:29139190"
/translation="MTAVFQSGKQHRVSEGGTQVLEKLDIATGTETFAEVLMTANG
EVKIGVPFDGQVKAEEVVAHGRGEKXIVKFRKKYRKKQGHROWFTDKITGIS
A"
gene
6703..7674
/genes="ispB"
/locus_tag="t3222"
6703..7674
/genes="ispB"
/locus_tag="t3222"
/notes="corresponds to STY3484 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="octaprenyl-diphosphate synthase"
/protein_id="AA070758.1"
/db_xref="GI:29139191"
/translation="MNLEKINELTAQDMAGVNAVITLQNLSDVQLINOLGYVYISGGG
KIRPMIAVLAARVYQGNVHTIAALIEPTHTLHDDVDVDESDWRGKATANAA
FGNAASVLVDFTYEAFOQMTSLGSLKVLVWSEAVNVIAEGEVLQLMVNDPDETE
ENMYRIVSKTARLEAAACQSGILAGCTPEQEGKQDIQRYGLVGFATFOLDLDISA
DEHGLKGVQDNLNCKPTLPULHARHGTPQESAMVITAEFGNGRHLHSPVLEANT
TCSLEWTKQAREEADKAISALQILPDTFWREALIGLAHIAVQRDR"
7907..8194
/genes="nlp"
/locus_tag="t3223"
7907..8194
/genes="nlp"
/locus_tag="t3223"
/notes="corresponds to STY3485 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="Ner-like regulatory protein"
/protein_id="AA070759.1"
/db_xref="GI:29139192"
/translation="MESKLIIDHPADIAGLRKKGTSMVAESRRNLSSSTIANLNR
1 73.7%; Score 289.8; DB 1; Length 299991;
Query Match
Best Local Similarity 85.2%; Pred. No. 1.5e-73;
Matches 335; Conservative 0; Mismatches 57; Indels 1; Gaps 1;
QY 1 TTCAACCTCAACGACGACGCGCGGATATCAATAAACAAGCAGCAATAGAGCGCGGTG 60
Db 65335 TTCAACCTCAACGACGACGCGCGGATATTAATAAACAAGCAGCAATAGAGCGCGGTAT 65394
QY 61 TGAACAAACCGCGGATCCGACAGAAAGTCTGTAGTACCTGTGGCACAAGTACCCGGCGCA 120
Db 65395 TGAACAGAGCGCGATGCCAGACAAATCATAGATCCCGTGGCACAAGTACCCAGCGCA 65454
QY 121 CCAGCCAGAAAATCCCGAGTATATAAAGCAATTCAAAAAATATCGTGGGAGGACAGCC 180
Db 65455 CCAGCCAGAAAATAGCAGTGACAAAAGCAATTCGAAAAATATCGTGGGAGGAGAC 65514
QY 181 CAACACACGATACGAGAACACACTCATGTTGACCTTGGTGTAAAGAGAGAGCGAGG 240
Db 65515 CAACACACGATACGAGGAAACACTCATGTTGACCTTGGTGTAAAGAGAGAGCGAGG 65574
QY 241 CGTTATTATTTTCAGCATCTGTGCGCGCAGAGAGGCGATGAAAGCGCGCGGAGAGCAA 300
Db 65575 CGATAGATTGTTAAGC-TACATCGCGCAGAGAGGCGCAAGAGCGCGCGGAGAGCGCA 65633
QY 301 CATTCTGTAGATTGATATTTAATATATATGCGTAACTGTGTATGCTGTATCTATATTAT 360
Db 65634 CATAGCTCTTATTCGAATTAAGTAGTATATTAACGTAATTTGTTATGCTGTATCTATATTAT 65693
```

```

QY 361 CTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
Db 65694 CTGATCTAAATCACTTTTAAGTCAGAGTGAATA 65726

RESULT 11
LOCUS
DEFINITION
Shigella flexneri 2a str. 301 section 301 of 412 of the complete
genome.
ACCESSION
AE015338 AE005674
VERSION
AE015338.1 GI:24053705
SOURCE
Shigella flexneri 2a str. 301
Shigella flexneri 2a str. 301
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Shigella.
REFERENCE
1 (bases 1 to 11744)
Jin, Q., Yuan, Z.H., Xu, J.G., Wang, Y., Shen, Y., Lu, W.C., Wang, J.H.,
Liu, H., Yang, J., Yang, F., Qu, D., Zhang, X.B., Zhang, J.Y., Yang, G.W.,
Wu, H.T., Dong, J., Sun, L.L., Xue, Y., Zhao, A.L., Gao, Y.S., Zhu, J.P.,
Kan, B., Chen, S.X., Yao, Z.J., He, B.K., Chen, R.S., Ma, D.L.,
Qiang, B.Q., Wen, Y.M., Hou, Y.D. and Yu, J.
Genome sequence of Shigella flexneri 2a: insights into
pathogenicity through comparison with genomes of Escherichia coli
K12 and O157
Nucleic Acids Res. 30 (20), 4432-4441 (2002)
12384590
REFERENCE
2 (bases 1 to 11744)
Jin, Q., Shen, Y., Wang, J.H., Liu, H., Yang, J., Yang, F., Zhang, X.B.,
Zhang, J.Y., Yang, G.W., Wu, H.T., Dong, J., Sun, L.L., Xue, Y.,
Zhao, A.L., Gao, Y.S., Zhu, J.P., Chen, S.X., Yao, Z.J., Wang, Y.,
Lu, W.C., Qiang, B.Q., Wen, Y.M. and Hou, Y.D.
Direct Submission
Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry
of Public Health, 100 YingXin Jie, XuanWu Qu, Beijing 100052, P.R.
China
FEATURES
Location/Qualifiers
1..11744
/organism="Shigella flexneri 2a str. 301"
/mol_type="genomic DNA"
/strain="301"
/serotype="2a"
/db_xref="taxon:198214"
/complement(228..620)
/genes="rpsI"
/locus_tag="SF3270"
/complement(228..620)
/genes="rpsI"
/locus_tag="SF3270"
/notes="Residues 1 to 131 of 131 are 99 pct identical to
residues 1 to 131 of a 131 aa protein from Escherichia
coli K12 gi: 1789625"
/codon_start=1
/transl_table=11
/product="30S ribosomal subunit protein S9"
/protein_id="AA04734.1"
/db_xref="GI:24053706"
/translation="MAENQYGTGRKSSAABVFIPKNGKIVINORSLEQVFGRETA
RMVROPMLVNVKLDYITVGGISGQAGAKHGITRALMEYDESLSELKAG
FVTRDARQVERKXVGLRKARRPQFSKR"
gene
/complement(636..1139)
/genes="rplM"
/locus_tag="SF3271"
/complement(636..1139)
/genes="rplM"
/locus_tag="SF3271"
/notes="Residues 26 to 167 of 167 are 100 pct identical to
residues 1 to 142 of a 142 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289799.1"
/codon_start=1
/transl_table=11
```


Db	11579	TGACACCGCGGATGCCACACAAAGTCGTAGATACCTGTGGCAACAAGTACCCGCGCA	11638	CDS	complement(1285..2730) /gene="tldb" /locus_tag="SF3283" /note="Residues 1 to 481 of 481 are 100 pct identical to residues 1 to 481 of a 481 aa protein from Escherichia coli K12 ref: NP_417711.1" /codon_start=1 /transl_table=1 /product="suppresses inhibitory activity of CsrA" /protein_id="AAN44747.1" /db_xref="GI:24053721"
Qy	121	CCAGCCAGAAAATCGCCAGTGATATAAAGCAATCAAAAATATCGGTGGGAAGCAGCC	180		/translation="MSNLVSEQLLAANGKHQDLFAILQQLAERRLDYGDLYPQSSY HESVLEDRILKDSYNIQGVRAISGKTFAYADQISLLALEQAQAARTIVR SGDKQVTLGAHVSPLYTSPVPLQSMREKLDILRRVKVAREADRVQEVNTASIS GVLEILVATDGTLAADVRPLVLSVLEEDGKRGASGGGSGFGEFFLADLD GSEVADAWAKEAVRMALVNLVSAAPAGTVPVLGACGPGVLLHEAVGHLEGDGNER GTSVSGVGELVASLCTVVDDGTWVDRGSVAIDDEGTPGQYVNLLENGILKGYNQ DKNARLMGMPGNGRRSEYAHLPMPRTNTYMLPKSTPQEIIESEYIAPNFG GGQVDTISGKFPVSTSEAYLIENGKVTKPVKGATLIGSGIETWQIQISWNGDKLQNG VVGCGEGQSLPVGVGQPTLTKVDNLTVGGTA"
Db	11639	CCAGCCAGAAAATCGCCAGTGATATAAAGCAATCAAAAATATCGGTGGGAAGCAGCC	11698		/gene="yhdRp" /locus_tag="SF3284" /note="Residues 1 to 1266 of 1266 are 99 pct identical to residues 1 to 1266 of a 1266 aa protein from Escherichia coli K12 ref: NP_417712.1" /codon_start=1 /transl_table=1 /product="Uncharacterized membrane protein" /protein_id="AAN44748.1" /db_xref="GI:24053722"
Qy	181	CAAAACCCAGATAACGGGAACAGACTCATGTGACCTGGTGTGT	226	gene	/translation="MRRPGILLTGAAVLVIAALLVSGLRILALPHAWRPEILNKI ESATGMPEVASQLSAQWNGPGLTEAHDIRAEIKDGGERSVKRVTLALDVQSLHMR WORDELTFWOLRFRTPITSGGSDSLASHISDLRFQFCHFDLSDSVSLTPSG ORALELAPQLTWLNDPRRRAEGLVSLISITGQGVQVNRMDLRDEGLSNGRWLQ ADIDILKPLGKWNQDIALETAQFSGEMTIDKGDVTGGVNLGQGLSAGWGEKQT HTLVNLTATHITRENPGWQFSIPDTRITMDGKWPSPGALTAWIPEQDVYGGDKRS DELRIRASINLEAGLEGIRPLAAKSPALGDVWRSTQPSGKINTALDIPLOAADKTR FQWSDSLAWKQWKLPGAEHFSGLTSGSVENGLLTASMKQAMPVETVFRAPLEIAD QOATISWNLKQGLDGRNIDVAKAVAHAGGRVYLOPANDPEWILGILAGISTDDGS QAWRYPENLWKGDLVDYLSGATOGGEADNATVYGVNQLPEYKNGHGEVPLR NAKFAEQPWALTNLDIELDFINDGLWKTGDNVGLGVASNLTAIVIDYSEKELLI DADIKGFGKAVGPFDETLKQSLGATLQELQDGDVNRALHLDIPLNELVYAKGEV TLRNSLFTKPLDLSILKSGKFSFNGDLQSEPTASFNQPLNVDFSTKEGAKAYQ VAVNLNGWQPAKTGVLPEAVNEALSGSVAMDKGVGIDLPYHAGATYVNLGDLNNV SSHLPSPKAPAGEPLAVNVKVDNLSNLPETLQAGADNHNFSRWLGLQKLTLDRAIW AADSKTLPPLPEOSGVELNMPMNGAEWLAFKGAESVGAASFPQHTLTPTMLS LGNQWNLSTVSQPTANGTVAOGRINATLANNAPLANIKYLYNPSVAKTR GDSPTSPPTTERINFRWPDQICRTECWGQKFGKGRIDSITISGDTLITLNGLI DTGSRLTADGEWNNPNERISLKGKRGQKIDAAAFFGVTTPIRQSFNVYDHL WRKAPWQDEATINGIHTOLGKEITEINTHAGQLLLSLVDALMKRLRFDRTF GEGFYDSIRSTAWIKDGMHTDDTLVDGLEADIAMKGSVNLVRDLNMEAVVAPIS ATVGAFAAFVNPVGAFAAASKVLGPLWSKVSILRYHISGLDDEPQINEVLRQPRK EKAQ"
Db	11699	CAAAACCCAGATAACGGGAACAGACTCATGTGACCTGGTGTGT	11744		complement(6754..8241) /gene="cafa" /locus_tag="SF3285" /note="Residues 1 to 495 of 495 are 99 pct identical to residues 1 to 495 of a 495 aa protein from Escherichia coli O157:H7 EDL933 ref: NP_289814.1" /codon_start=1 /transl_table=1 /product="bundles of cytoplasmic filaments" /protein_id="AAN44749.1" /db_xref="GI:24053723"
RESULT 12				gene	/translation="MRKGINNTAEILLVNTVPSERVAVYDGGILQEIHLIERBARGIV GNLYKGRSRVLPFGMAAFVIGLDKAAFLHASDIMPTECVAGSEQKQFTVRDISEL
AE015339					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
gene					
CDS					
gene					

VRQGDLMVQVVKPLGTGKARLTDTITPLSRVLFVFMGASHGVQSRIESESERERL
KVAAYCDEQGGFIIRTAAGGVGAELASDAAYLKRVTKVMERKKRPOTRYQLYGE
LALAQRVLDFADELDIRVDSRLTYEALLFTSEYIPENTSKLHVTPGPIFDLF
DVENIQALERKVELKSGVLLIDQTEAMTVDINTGAFVGRNLDITPNTWIEAT
QAIARQLRNLGGIIIDIFDMNEDHRRVRLSHLEQALSADRVKSVNGFSALGV
ENRKRTRRESIHVVJCNBCTHGRTGTVKTEYELMREIIRVHRAIDSRFLVTA
SPAVAAKGEESHSLAEVIFVGRQVQVQIPLYNQOFQVVMW
/gene="yhde"
/locus_tag="SF3286"
complement(8213..8806)
/gene="yhde"
/locus_tag="SF3286"
complement(8213..8806)
/note="Residues 1 to 197 of 197 are 97 pct identical to
residues 21 to 217 of a 217 aa protein from *Escherichia*
coli gb: AAAS8051.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA04750.1"
/db_xref="GI:24053724"
/translation="MTSLVLAGSPRRQELLAQLGVTFEIVTGIEQRPQBSAQY
VRLAREKAGVAGTADLPVLGADIIVILNGEVLKPRDAEHAQVRLKLSGTHQ
VMTAVALDSOHLCLVVDVTFRTLDEIDAGVVASGEPLDKAGAYGIQGLGCFV
RKINGSYHAVUGLPLVETVELLSNPNALREKDKDG"
complement(8815..9303)
/gene="mreD"
/locus_tag="SF3287"
complement(8815..9303)
/gene="mreD"
/locus_tag="SF3287"
/note="Residues 1 to 162 of 162 are 98 pct identical to
residues 1 to 162 of a 162 aa protein from *Escherichia*
coli O157:H7 EDL933 ref: NP_289816.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA04751.1"
/db_xref="GI:24053725"
/translation="WASYRSQGRVWISFLALLQLQMPEDNLIIVFRPNWLLILL
YMLALPHRVNVTGTFVMAIGLDSGLMGLVRLVMSIIAYLVALKYOLFRNLALWQ
QALVVMLLSLVDIIVFAEFLVNVSEFVSWVNGVLPWIFLLMRKVRQQA
VQ"
complement(9303..10406)
/gene="mreC"
/locus_tag="SF3288"
complement(9303..10406)
/gene="mreC"
/locus_tag="SF3288"
/note="Residues 1 to 367 of 367 are 99 pct identical to
residues 1 to 367 of a 367 aa protein from *Escherichia*
coli K12 ref: NP_417716.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA04752.1"
/db_xref="GI:24053726"
/translation="MKPIFSRPSLQIRLILAVLALGIIIDSLRGLTFSQIRTYMDT
AVSPFFVSNAPRELLDGI SOTLASDQLENLRLRQELLKNSLLMLGQYKQENA
RLRELGLSPRDEQKQVTCVISTVNDPYSQDVQVVDKGSVNGVYEGQVSPKGVVQ
VVAVALTRVLICDHALPIQVLRNDIRVIAAGNCTDQLQELHPLPNTDIRUGD
VLVTGLGRFPPEYVAVSVYKLTQRAYTVIQAQRTAGLQRLYLLLLMGDRNG
ANPMTEPVRHVRANERLMQVQLSPDMGPKLPEPATGIAQTPPOQPATGNAATA
PVAPTQPAANRSPQATPPOSQAQPPAPPGG"
complement(10472..11590)
/gene="mreB"
/locus_tag="SF3289"
complement(10472..11590)
/gene="mreB"
/locus_tag="SF3289"
Query Match 57.4%; Score 225.4; DB 1; Length 11734;
Best Local Similarity 99.6%; Pred. No. 9e-55;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
167 TGGGAGGAGAGCCCAACACGATACCGGAAACAGACTCATGTTGACCTTGGTTGT 226

1 TGGGAGGAGAGCCCAACACGATACCGGAAACAGACTCATGTTGACCTTGGTTGT 60
227 AAAGAGAGAGCAGCGCCTTATTATTTTCAGCATCTGTGCGCGAGAGAGGGCATGGAAG 286
61 ATAGAGAGAGCAGCGCTTATTATTTTCAGCATCTGTGCGCGAGAGAGGGCATGGAAG 120
287 CCGGGGAGAGCAACATTGCTGTAGATTGATATTTAATATATATAGCTAATGCTTATGCT 346
121 CCGGGGAGAGCAACATTGCTGTAGATTGATATTTAATATATATAGCTAATGCTTATGCT 180
347 GTTATCTATTATGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
181 GTTATCTATTATGATCTAAATCACTTTTAAGTCAGAGTGAATA 227
RESULT 13
AR389738/c
LOCUS AR389738 258 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6467 from patent US 6610836.
ACCESSION AR389738
VERSION AR389738.1 GI:40099919
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 258)
AUTHORS Breton, G.L. and Osborne, M.
TITLE Nucleic acid amino acid sequences relating to *Klebsiella pneumoniae*
for diagnostics and therapeutics
JOURNAL Patent: US 6610836-A 6467 26-AUG-2003;
FEATURES
source 1..258
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 37.3%; Score 146.6; DB 6; Length 258;
Best Local Similarity 78.3%; Pred. No. 8.5e-32;
Matches 176; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
8 TCAACGAAACAGTCGGATATCAATAAACAACAGCAGCAGTACGCGGTGTTGAACA 67
258 TCAACGAAACAGTCGGATATCAATAAACAACAGCAGCAGTACGCGGTGTTGAACA 199
68 CGCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACAGCCA 127
198 TCGTGGTGCCAGACAAATCGTAATGCCGGTCGGAACCCAGCAGTCGGTGCCAGCCA 139
128 GAAATCGCAGTGATAAAGCAATTCAAAAATATCGTGGGAGAGAGAGCCCAACAC 187
138 AAAGATAGCCAGCAGCAAAAGCACTCGAAGAGATCGGCGGAAACGATAAGCCGAAAT 79
188 CACGATACCGGAAACAGACTCATGTTGACCTTGGTTGTAAGAG 232
78 CACGATACCGGAAACAGACTCATGATGACCTTGAATAGAAGCG 34
RESULT 14
AB080601/c
LOCUS AB080601 593 bp DNA linear BCT 01-MAR-2002
DEFINITION *Serratia marcescens* yhcR gene for hypothetical protein, complete
cds.
ACCESSION AB080601
VERSION AB080601.1 GI:19032360
KEYWORDS
SOURCE *Serratia marcescens*
ORGANISM *Serratia marcescens*
REFERENCE 1
AUTHORS Baba, M., Midorikawa, Y., Nakagawa, Y., Fujita, M. and Matsuyama, T.
TITLE *Serratia marcescens* and *Escherichia coli* genes controlling

temperature-dependent production of structurally unrelated secondary metabolites such as prodigiosin and serrawettin unpublished

2 (bases 1 to 593)

Nakagawa, Y.; Baba, M. and Matsuyama, T.
Direct Submission
Submitted (27-FEB-2002) Xoji Nakagawa, Niigata University, Faculty of Agriculture; 8050, Karashi 2-no-cho, Niigata City, Niigata 950-2181, Japan (E-mail:nakagawaagr.niigata-u.ac.jp, Tel:81-25-262-6360(ex.6360), Fax:81-25-262-6360) Location/Qualifiers

FEATURES
source

```

source
1..593
/organism="Serratia marcescens"
/mol_type="genomic DNA"
/db_xref="taxon:615"
301..504
/gene="yhcr"
301..504
/gene="yhcr"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="BAB85653.1"
/db_xref="GI:19032361"
/translation="MSLFLPMVIFGLSFPPV
HPALFATALYCCFLVLSICFLV"

```

ORIGIN

Query Match	32.0%;	Score 125.8;	DB 1;	Length 593;
Best Local Similarity	75.1%;	Pred. No. 1.11e-25;		
Matches 157;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;
QY	3	CAACCTCAACGGAACAGTCGCGATATCAAATAAAACAACGACCAATAGAGCGCGGTG	52	
Db	509	CGACCTCAACAAAAGACATGAATCAGATAAAACAACGACCAATACGCCAGTGTTA	450	
QY	63	AACAAGCGCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAGTACCCGGCGCAC	122	
Db	449	AACAGCGCGGATGCCAGACAAATCGTAATCCCGTCGCGTGCAGAGGCGGCGCAAG	390	
QY	123	AGCCAGAAAAATCGCCAGTGATAAAGCAANTTCAAAAAATATCGTGGGAAGACAGCCCA	182	
Db	389	AGAAAAACAGCGGCACCGGCACACGACGCTCGAAAAACCCGGCGAAACACACAGTC	330	
QY	183	AACACACGATACCGGAAACACAGCTCAT	211	
Db	329	AAGATGACCAATAACCGGAAGCAACTCAT	301	

RESULT 15

AE013617 LOCUS 10553 bp DNA linear BCT 26-JUL-2002
DEFINITION *Yersinia pestis* KIM section 17 of 415 of the complete genome.
ACCESSION AE013617 AE009592
VERSION AE013617.1 GI:21956840

KEYWORDS

SOURCE	ORGANISM
Yersinia pestis KIM	
Yersinia pestis KIM	
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Yersinia.	

REFERENCE

AUTHORS

Deng, W., Burland, V., Plunkett, G. III, Boutin, A., Mayhew, G.F.,
Liss, P., Perna, N.T., Rose, D.J., Mau, B., Zhou, S., Schwartz, D.C.,
Fetherston, J.D., Lindler, L.E., Brubaker, R.R., Plana, G.V.,
Straley, S.C., McDonough, K.A., Nilles, M.L., Matson, J.S.

TITLE

1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500
 1501
 1502
 1503
 1504
 1505
 1506
 1507
 1508
 1509
 1510
 1511
 1512
 1513
 1514
 1515
 1516
 1517
 1518
 1519
 1520
 1521
 1522
 1523
 1524
 1525
 1526
 1527
 1528
 1529
 1530
 1531
 1532
 1533
 1534
 1535
 1536
 1537
 1538
 1539
 1540
 1541
 1542
 1543
 1544
 1545
 1546
 1547
 1548
 1549
 1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560
 1561
 1562
 1563
 1564
 1565
 1566
 1567
 1568
 1569
 1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579
 1580
 1581
 1582
 1583
 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640

Straley, S.C., McDonough, K.A., Nilles, M.L., Matson, J.S.,
Blattner, F.R. and Perry, R.D.
Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA

FEATURES

source

```

"organism"="Yersinia pestis KIM"
/mol_type="Genomic DNA"
/strain="KIM"
/db_xref="taxon:187410"
complement(129..1469)
/gene="pmba"
/note="synonym: y0171"
complement(129..1469)
/gene="pmba"
/function="phenotype; proteins - translation and modification"

```

gene

```
/gene="pmbA"
```

CDS

```

Complement=23. 114937
Gene="PmbA"
/function="phenotype; proteins - translation and
modification"
/notes="maturation of antibiotic MccA17, see tid 9
residues 1 to 446 of 446 are 82.73 pct identical
residues 5 to 450 of 450 from E. coli K12 : B423"
/codon_start=1
/transl_table=11
/product="PmbA/tldD family protein"
/protein_id="AA083765.1"
/db_xref="GI:12195841"

```

gene

```

-044.1220
/gene="v0172"

```

CDS

```

/genes="y0172"
/function="phenotype; Not classified"
/notes="residues 3 to 183 of 184 are 78.02 pct identical to
residues 1 to 182 of 183 from E. coli K12 : B4234"
/codon_start=1
/transl_table=11

```

gene

complete (2010, 2011)
/gene="v0173"

CDS

```

/compname="Y0173"
/gene="Y0173"
/notes="residues 6 to 87 of 93 are 53.01 pct identical to
residues 6 to 88 of 90 from E. coli K12 : B3239"
/codon_start=1

```

gene

```
comp = compile(
    /gene="y0174"
```

CDS

```

/notes= residues 1 to 156 of 157 are 47.43 pct identical to
/notes= residues 1 to 148 of 149 from GenPept : >GBAA86441.1
(M14442) barnase (RNase) precursor [Bacillus
amyloliquefaciens]"
/function="enzyme; macromolecule degradation: Degradation
of RNA"
/genes="y017a"

```

gene
complement (6956..7891)
/gene="y0178"
CDS
complement (6956..7891)
/gene="y0178"
/function="putative membrane"
/note="residues 1 to 311 of 311 are 72.66 pct identical to
residues 1 to 310 from E. coli K12 : B3241;
residues 1 to 311 are 100.00 pct identical to
residues 1 to 311 of 311 from GenPept : >emb|CAC93154.1|
(AJ414158) putative HlyD family secretion protein
[versinia pestis]"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="AA083772.1"
/db_xref="GI:21956848"
/translation="MSTFSLKIRVGVTVLVVAVIAIFNVWAFVETSPWTRDAKFT
ADVVAIPDVSGLLTEVPKQNLVQKQILFVIDPQYQQAALAEADVAYVOTLAA
EKQRESSRHRIGIQALESQEDIDQSNVLTQVHQAKAIAVRLDLRLTTRAP
ABGWNTLVNHAGEFINRGATAVALVKDTPFVILAYLEETKLEGVKPGFRAETIPGS
NRILHGTVDISIGAVTINSSSSADSKGLATIDNNLEWRLAQRPVVKIHLDSDDQOYLY
PAGTTATVITGTNDRDPHQASPMTKLMHRLREFG"
complement (7899..8102)
/gene="y0179"
complement (7899..8102)
/gene="y0179"
/note="residues 1 to 66 of 67 are 81.81 pct identical to
residues 24 to 89 of 90 from E. coli K12 : B3242"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 29.9%; Score 117.4; DB 1; Length 10553;
Best Local Similarity 72.9%; Pred. No. 4e-23;
Matches 151; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 5 ACCTCAACGACAGTCGCGATATCAATAAACAAGCAGCAGATAGCGCGGTGTAA 64
Db 7896 ACCTCAGGAAAAAGACGGATGTTAAATAAACAAGCAGCAGATAGCGCGGTATAA 7955
Qy 65 CAACGCGCGATGCCAGCAAAAGTCGTAGATACCTGTGGCACAAGTACCGGCGCACAG 124
Db 7956 AAGGCGCGGTGCCAACAACCTCGTAATACCGGTGGTGCAGGATGCGGCGCACAC 8015
Qy 125 CCAGAAATCGCAGTGATTAAGCAATTCAAAAAATATCGGTGGAGGACAGCCCCAA 184
Db 8016 AAAAAACAAGCCCAATGAAATTAGCAGCTCAAGAAATATCGGTGGAGGAGGCCAA 8075
Qy 185 CACCAGCATACCGGAAACAGACTCAT 211
Db 8076 AATCACCATCACCAGGACAACTCAT 8102
Search completed: December 19, 2004, 01:09:55
Job time : 1930.01 secs

gene
complement (3264..3602)
/gene="y0175"
CDS
complement (3264..3602)
/gene="y0175"
/function="unknown"
/note="residues 1 to 112 of 112 are 59.82 pct identical to
residues 374 to 485 of 486 from GenPept : >gb|AA044216.1|
(AE009270) succinate semialdehyde dehydrogenase
[Agrobacterium tumefaciens str. C58 (U. Washington)]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AA083769.1"
/db_xref="GI:21956848"
/translation="MIADANEQMKVASEETFGPLAACRPFSTEAEVIOQRANTPFGILA
AYFTQNLQVRVSEALLESIGMIGVNECSVSTELAPFGGVKESGLREGSVLGLDEFM
EVKTLHLGNL"
complement (3599..4732)
/gene="gabD"
complement (3599..4732)
/gene="gabD"
/function="enzyme; central intermediary metabolism: Pool,
multipurpose conversions"
/note="NADP-dependent activity; residues 23 to 366 of 377
are 56.06 pct identical to residues 15 to 360 of 482 from
E. coli K12 : B2661"
/codon_start=1
/transl_table=11
/product="succinate-semialdehyde dehydrogenase"
/protein_id="AA083770.1"
/db_xref="GI:21956846"
/translation="WPAONPLOHQDHGIVHIGVFGKWSQAQDFVVPATGEL
VARVAKSQFTEAIAKASEAFPAWRKTPAKORAEILQRYVLIIMEHQSLAELMVS
EOKPKUEALVEVAASFIQFSEQAKRANGELIIPSAKEGARILATREPVGVAAT
PWFPLMLTRKLGALPAGTGLIKPANNTPLSFALLALAEQAGVPGVINGVVD
THAISDAIMASSDVRIKISFTSTNGTKLMRNAATMKRI SNELGNNAPYIVFDDADL
EAAVAGAMACKFRNAQGVQVNVRFYQDAIYDEFVRLAAAEVKLVKVGNGMDKDVNN
GPLINLAGLEKVEEHVDALKEGRLLAGIAMRWAVISSOR"
complement (4999..6954)
/gene="y0177"
complement (4999..6954)
/gene="y0177"
/note="residues 9 to 650 of 651 are 64.64 pct identical to
residues 11 to 652 of 655 from E. coli K12 : B3240;
residues 9 to 650 of 651 are 64.79 pct identical to
residues 11 to 652 of 655 from GenPept :
>gb|AAG59368.1|AE005551.1 (AE005551) orf, hypothetical
protein [Escherichia coli O157:H7 EDL933]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA083771.1"
/db_xref="GI:21956847"
/translation="MTHPSFIRLRFAPKLSFAVAALFLGHLQLETPRWSVLTAAIV
SAGPAFAAGPEPSGAIRKGMRLIGTFIGGILVILVITRAPVLTMLCCLWAG
ICTWISLVREVSAYFAGLGYTALIIVTTGETPLTPQFAVERCSEIVIGVCANM
ADLLFSPRIKODIDRLVDKLVQYRLQLCIQPAKSEIDRAWNELVKNITSLNGM
RSLVLMWESRQRCNRLQVLTESLALITQACETVYVNSNHPVISAELKMLSEPA
OTPAETHQMKLROFIAASHGEAIPHTISWVGAATRYLLSKIGTNSINQVEED
ILAGDAPVPLSAGEHAMNGRTGATAIGLFWLTGWTSGAGCMVLAFTSLA
MRTPNRRALDPLVGVITLALPICALYFMFIIPSTQOSMLLICI SLGLVLAFTIGIEVQ
KRRIGSLTASTINIVLSNPMIFNVRFDSALGQIVGCVFSLVILLITPDNAKDR
TGRLLNRFYSVSAVLTNKTGRGENHLPALYQNLNQLMFPFADIDKRLALTLII
AHQRLNRTEIPVNAELSAFHKQIRSTAERVI TVNNDQKRYYPARLLQELDQYQKLV

GenCore version 5.1.6
Copyright' (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:46:30 ; Search time 257.841 Seconds
(without alignments)
8001.134 Million cell updates/sec

Title: US-10-759-889-3

Perfect score:

Sequence: 1 ttcaacctcaaacgaacagt.....cttttaagtcagaagtqaata 393

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs; 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB sea length. n

Maximum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : N_Geneseq_23Sep04.*
1: 1: geneseqn1980s.*
2: 2: geneseqn1990s.*
3: 3: geneseqn2000s.*
4: 4: geneseqn2001as.*
5: 5: geneseqn2001bs.*
6: 6: geneseqn2002as.*
7: 7: geneseqn2002bs.*
8: 8: geneseqn2003as.*
9: 9: geneseqn2003bs.*
10: 10: geneseqn2003cs.*
11: 11: geneseqn2003ds.*
12: 12: geneseqn2004s.*
```

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	146.6	37.3	258	11	ABD00692	Klebsiell
C 2	100	25.4	100	8	ABD69190	E. coli K
C 3	100	25.4	100	8	ACD69191	E. coli K
C 4	100	25.4	100	8	ACD69192	E. coli K
C 5	38.4	9.8	1181	2	AAV12742	E. coli K
C 6	37.2	9.5	1035	11	ACH95383	RBE1 tran
C 7	37	9.4	35000	6	ABT12817	ACH95383 Klebsiell
C 8	36.8	9.4	13503	6	ABL34071	Human imm
C 9	36.6	9.3	397	4	ABL16233	Human imm
C 10	36.6	9.3	397	4	ABA58948	Probe #61
C 11	36.6	9.3	397	4	AAI38665	Human foe
C 12	36.6	9.3	397	4	ABA27808	Human foe
C 13	36.6	9.3	397	4	AAK32852	Probe #73
C 14	36.6	9.3	397	4	AAK07110	Human bon
C 15	36.6	9.3	397	4	ABS32578	Human bra
C 16	36.6	9.3	397	6	ABS07656	Human liv
C 17	35.6	9.1	711	2	AAAT67444	Human gen
C 18	35.6	9.1	723	2	AAAT67946	H. pylori
C 19	35	8.9	110000	5	AAI613713	H. pylori
C 20	34.8	8.9	4223	4	ABL17172	Continuance (3 of
C 21	34.8	8.8	14041	4	AAH48024	AB117172 Drosophil
C 22	34.8	8.8	14041	4	AAH48024	Internal

Abg67094 Human ang
Continuation (12 o
Abk15056 A. thalia
Adc47192 DNA sequ
Abg39467 Oligonuc
Abg39466 Oligonuc
Adm67052 Murine ad
Adm62896 A. gossyp
Aav52303 Streptoc
Continuation (8 of
Aac39627 Arabidops
Abi013322 Drosophi
Abk79579 Bacillus
Abi17948 Human bre
Aal10931 Human ova
Adc34366 Human cod
Abg465886 Oligonuc
Abg46587 Oligonuc
Aca33079 Prokaryot
Abz10017 Haematopo
Abz10163 Haematopo
Ade48205 Human lym
Ade44129 Human lym
Abi14476 Drosophi

ALIGNMENTS

RESULT 1

ABD00692/c

ID ABD00692 standard: DNA: 258 bp.

AC ABD00692;

XX

DT 29-JUL-2004 (first entry)

XX
DE Klebsiella pneumoniae polynucleotide seqid 6467.XX
XX

KW Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae

XX

OS Klebsiella pr

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PN
yy
US6610836-B1

XX
PD
26-AUG-2003

ED
XX
26-AUG-2003.

27-JAN-2000: 2000US-00489039

FF
XX
27-04N-2000; 2000US-00489039.

29-JAN-1999: 99JIS-

[illegible]

PA (GENO-) GENOME THERA

XX
XX
CONTENTS (continued)
XX

PI Breton GL, Osborne M;

WPI: 2003-895346/82.
P-PSDB; ABO67121.

New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.

Disclosure: SEQ ID NO 6467; 932pp; English.

The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This sequence encodes a Klebsiella pneumoniae polypeptide of the invention

Sequence 258 BP; 40 A; 66 C; 53 G; 99 T; 0 U; 0 Other;

Sequence 258 BP; 40 A; 66 C; 53 G; 99 T; 0 U; 0 Other;

Query Match 37.3%; Score 146.6; DB 11; Length 258;
Best Local Similarity 78.2%; Pred. No. 9.5e-36; Mismatches 49; Indels 0; Gaps 0;
Matches 176; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
XX Sequence 100 BP; 16 A; 24 C; 23 G; 37 T; 0 U; 0 Other;
QY 8 TCAACGACAGTCGCGATATCAATAAAACAGCAGCAATAGAGCGCGGTGTGAACAA 67
Db 258 TCAACGACAGTCGCGATATCAATAAAACAGCAGCAATAGAGCGCGGTGTGAACAG 199
QY 69 CGCGGATGCCAGACAAAGTGTAGATACCTGTGGCAAGTACCGCGCCAGCCCA 127
Db 198 TGCTGGTGCCAGACAAATCGTAAATCGGTCGGTCCAGCAGTCCGTCAGCCAGCA 139
QY 128 GAAATCGCAGTATATAAAGCAATCAAAAAATATCGTGGGAGGAGCAGCCCAACAC 187
Db 138 AAGATAGCCAGCAGCAAAAGCACTCGAAGAGATCGCGGAAACGATAGCGGAAAT 79
QY 188 CAGGATACGGGAAACAGACTCATGTGACCTGTGTTGTAAGAG 232
Db 78 CAGGATACGGGAAACAGACTCATGTGACCTGTGTAAGAGCG 34

RESULT 2
ACD69190/c
ID ACD69190 standard; DNA; 100 BP.
XX ACD69190;
XX ACD69190;
DT 18-SEP-2003 (first entry)
DE E. coli K12 MG1655 biochip probe SEQ ID 460.
XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX Escherichia coli.
XX EP1260592-A1.
XX 27-NOV-2002.
XX 17-MAY-2001; 2001EP-00112179.
XX 17-MAY-2001; 2001EP-00112179.
XX (MWGB-) MWG-BIOTECH AG.
XX Donner H, Drescher B, Huber A, Weber J;
XX WPI; 2003-241155/24.
XX Biochip containing probes complementary with open reading frames in
XX Escherichia coli K12, useful for detecting gene expression and expression
XX patterns.
XX Claim 3; Page 82; 2004pp; German.

This invention describes a novel biochip comprising probe spots, each containing many identical probes. The probes are nucleotide sequences of 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at least one includes a segment of at least 20 bases identical with, or complementary to, a segment of an open reading frame (orf) of Escherichia coli K12. The biochip is used for specific detection of gene expression in K12 and for determining the gene expression pattern, e.g. for diagnostic determination of which E. coli strains are present in the gut, and to determine the effects of e.g. growth media on gene expression. The biochip provides a comprehensive analysis of many different genes with a single genome, with simultaneous analysis of gene expression between K12 and its mutants or other E. coli strains in a single experiment. Apart from qualitative and quantitative information about gene expression, it also allows measurements of population densities for the various strains. The use of synthetic oligonucleotides for preparation of probes allows free variation in probe length and ensures high purity (and thus selectivity, shorter than probes prepared by polymerase chain reaction. ACD68731 to

CC reactivity and reproducibility); also synthetic probes are generally shorter than probes prepared by polymerase chain reaction. ACD68731 to ACD81540 represent oligonucleotide probes used with the biochip described in the invention
XX Sequence 100 BP; 16 A; 24 C; 23 G; 37 T; 0 U; 0 Other;
QY 8 TCAACGACAGTCGCGATATCAATAAAACAGCAGCAATAGAGCGCGGTGTGAACAA 67
Db 100 TCAACGACAGTCGCGATATCAATAAAACAGCAGCAATAGAGCGCGGTGTGAACAA 41
QY 68 CGCGGATGCCAGACAAAGTGTAGATACCTGTGGCACA 107
Db 40 CGCGGATGCCAGACAAAGTGTAGATACCTGTGGCACA 1

RESULT 3
ACD69191/c
ID ACD69191 standard; DNA; 100 BP.
XX ACD69191;
XX ACD69191;
DT 18-SEP-2003 (first entry)
DE E. coli K12 MG1655 biochip probe SEQ ID 461.
XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX Escherichia coli.
XX EP1260592-A1.
XX 27-NOV-2002.
XX 17-MAY-2001; 2001EP-00112179.
XX 17-MAY-2001; 2001EP-00112179.
XX (MWGB-) MWG-BIOTECH AG.
XX Donner H, Drescher B, Huber A, Weber J;
XX WPI; 2003-241155/24.
XX Biochip containing probes complementary with open reading frames in
XX Escherichia coli K12, useful for detecting gene expression and expression
XX patterns.
XX Claim 3; Page 82; 2004pp; German.

This invention describes a novel biochip comprising probe spots, each containing many identical probes. The probes are nucleotide sequences of 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at least one includes a segment of at least 20 bases identical with, or complementary to, a segment of an open reading frame (orf) of Escherichia coli K12. The biochip is used for specific detection of gene expression in K12 and for determining the gene expression pattern, e.g. for diagnostic determination of which E. coli strains are present in the gut, and to determine the effects of e.g. growth media on gene expression. The biochip provides a comprehensive analysis of many different genes with a single genome, with simultaneous analysis of gene expression between K12 and its mutants or other E. coli strains in a single experiment. Apart from qualitative and quantitative information about gene expression, it also allows measurements of population densities for the various strains. The use of synthetic oligonucleotides for preparation of probes allows free variation in probe length and ensures high purity (and thus selectivity, shorter than probes prepared by polymerase chain reaction. ACD68731 to

CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention
XX
SQ Sequence 100 BP; 13 A; 22 C; 27 G; 38 T; 0 U; 0 Other;
Query Match 25.4%; Score 100; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.9e-21;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 CTGTTGGCACAAGTACCGGGCCGACGACGAGCAAGAAATCGCAGTGAATAAGCAATTCA 155
DB 100 CTGTTGGCACAAGTACCGGGCCGACGACGAGCAAGAAATCGCAGTGAATAAGCAATTCA 41
QY 156 AAAAATATCGTGGGAGGACAGCCCAACACACAGATAA 195
DB 40 AAAAATATCGTGGGAGGACAGCCCAACACACAGATAA 1
RESULT 4
ACD69192/C
ID ACD69192 standard; DNA; 100 BP.
XX
AC ACD69192;
XX
DT 18-SEP-2003 (first entry)
XX
DE E. coli K12 MG1655 biochip probe SEQ ID 462.
XX
KW Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX
OS Escherichia coli.
XX
PN EP1260592-A1.
XX
PD 27-NOV-2002.
XX
PF 17-MAY-2001; 2001EP-00112179.
XX
PR 17-MAY-2001; 2001EP-00112179.
XX
PA (MWGB-) MWGB-BIOTECH AG.
XX
PI Donner H, Drescher B, Huber A, Weber J;
XX
DR WPI; 2003-241155/24.
XX
PT Biochip containing probes complementary with open reading frames in
PT Escherichia coli K12, useful for detecting gene expression and expression
PT patterns.
XX
PS Claim 3; Page 82; 2004pp; German.
XX
CC This invention describes a novel biochip comprising probe spots, each
CC containing many identical probes. The probes are nucleotide sequences of
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
CC least one includes a segment of at least 20 bases identical with, or
CC complementary to, a segment of an open reading frame (orf) of Escherichia
CC coli K12. The biochip is used for specific detection of gene expression
CC in K12 and for determining the gene expression pattern, e.g. for
CC diagnostic determination of which E. coli strains are present in the gut,
CC and to determine the effects of e.g. growth media on gene expression. The
CC biochip provides as comprehensive as possible detection of the K12
CC genome, with simultaneous analysis of many different genes with a single
CC device, and comparison of gene expression between K12 and its mutants or
CC other E. coli strains in a single experiment. Apart from qualitative and
CC quantitative information about gene expression, it also allows
CC measurements of population densities for the various strains. The use of
CC synthetic oligonucleotides for preparation of probes allows free
CC variation in probe length and ensures high purity (and thus selectivity,
CC reactivity and reproducibility); also synthetic probes are generally
CC shorter than probes prepared by polymerase chain reaction. ACD8731 to
CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention

XX
SQ Sequence 100 BP; 23 A; 26 C; 20 G; 31 T; 0 U; 0 Other;
Query Match 25.4%; Score 100; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.9e-21;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 161 TATCGTGGGAGGACAGCCCAACACACAGATAACGGGAAACAGACTCATGTTGACCTT 220
DB 100 TATCGTGGGAGGACAGCCCAACACACAGATAACGGGAAACAGACTCATGTTGACCTT 41
QY 221 GGTTGTAAGAGAGAGAGAGCGGCTTATTATTATTTTCAGCATCT 260
DB 40 GGTTGTAAGAGAGAGAGAGCGGCTTATTATTATTTTCAGCATCT 1
RESULT 5
AAV12742
ID AAV12742 standard; cDNA to mRNA; 1181 BP.
XX
AC AAV12742;
XX
DT 05-MAY-1998 (first entry)
XX
DE RBE1 transcription controlling factor #33 (RBEF1) coding sequence.
XX
KW RBE1; rice starch branching enzyme; transcription controlling factor; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 55..852
FT tag= a
XX
PN JPI0014575-A.
XX
PD 20-JAN-1998.
XX
PF 28-JUN-1996; 96JP-00169246.
XX
PR 28-JUN-1996; 96JP-00169246.
XX
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA (MITK) MITSUI TOATSU CHEM INC.
XX
DR WPI; 1998-138237/13.
XX
P-PSDB; AAW41152.
XX
PT Transcription controlling factor in rice - comprises starch branching
PT enzyme gene, allows rice to be improved in quality and quantity.
XX
PS Claim 3; Page 6-7; 11pp; Japanese.
XX
CC This sequence represents the DNA encoding a transcription controlling
CC factor of the invention, designated #33 (RBEF1). The transcription
CC controlling factor in rice promotes transcription of the RBE1 gene. Use
CC of the transcription factor enables rice to be improved in quality and
CC quantity
XX
SQ Sequence 1181 BP; 315 A; 257 C; 337 G; 272 T; 0 U; 0 Other;
Query Match 9.8%; Score 38.4; DB 2; Length 1181;
Best Local Similarity 56.2%; Pred. No. 0.21;
Matches 72; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 126 CAGAAATCGCCAGTGTATAAAGCAATTCAAAATAATATCGTGGGAGGACAGCCCAAC 185
DB 778 CACAAGACACTAATAATACACACACGGGAACTATGTTAGTGGCAGCAACACAGAG 837
QY 186 ACCACGATACCGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGAGCGGTGA 245
DB 838 GCTAACTCTAGGTGAGAGAGATCATGAGCAAGTATGAGATGAGATGAGTGTTC 897

QY 246 TTATTTC 253
 Db 898 TTCTTATC 905

RESULT 6

ACH95383/c
 ID ACH95383 standard; DNA; 1035 BP.

XX ACH95383;
 AC ACH95383;

XX 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polynucleotide seqid 1178.

XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.
 XX Klebsiella pneumoniae.

OS US6610836-BL.
 PN 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;
 PI WPI; 2003-895346/82.
 DR P-PSDB; ABO61832.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX Disclosure; SEQ ID NO 1178; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
 CC pneumoniae polypeptide of the invention

XX Sequence 1035 BP; 151 A; 330 C; 312 G; 242 T; 0 U; 0 Other;

Query Match 9.5%; Score 37.2; DB 11; Length 1035;
 Best Local Similarity 50.0%; Pred. No. 0.48;
 Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 63 AACACGCGCGATGCCAGCAAGTGGTAGATACCTGTGGCACAAGTACCGCGGCACC 122
 Db 731 AACGCTCGCGCGCGCGACACATAGTGGTAGTGGCGGTACCGAGAAATCCGCGGCACC 672
 QY 123 AGCCAGAAATCCGCGATGATAAAGCAATTCAGGTTGGGGAAGACAGCCCA 182
 Db 671 AGCCGACGAGGTCATCCAGCAAGAGACGCGTATTCAGCAGCGCCACGAGAGGCC 612
 QY 183 AACACACGATACCGGAAACAGACTCATGTTGACCTGGTTGTAAGAGAGAGAGCGC 242
 Db 611 GGCTGCACGCGCTGGAGAGCAGGATATCAGAAACAGCGTTGGCATCCGCGACGACT 552
 QY 243 TTATTA 248
 Db 551 TCATA 546

RESULT 7

ABT12817

ID ABT12817 standard; DNA; 35000 BP.

XX ABT12817;
 AC ABT12817;

XX 16-JAN-2003 (first entry)

XX Human RECQL gene-related nucleotide #1.

XX Human; antisense therapy; ds; RECQL; hyperproliferative disorder; cancer;
 KW premature ageing; infection; inflammation; tumour formation.
 XX Homo sapiens.

OS WO200268590-A2.
 PN 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005225.

XX 23-FEB-2001; 2001US-00793807.

XX (ISIS-) ISIS PHARM INC.

XX Ward DT, Watt AT;
 PI WPI; 2002-750415/81.

XX New antisense oligonucleotides targeted to a nucleic acid encoding RECQL,
 PT useful for modulating the expression of RECQL protein, or for treating a
 PT disease or condition associated with the expression of RECQL, e.g.
 PT cancer.

XX Example 15; Page 101-121; 138pp; English.

XX The invention comprises antisense oligonucleotides which inhibit
 CC expression of the human RECQL gene. The antisense oligonucleotides of the
 CC invention are useful for modulating the expression of RECQL protein and
 CC in treating hyperproliferative disorders (e.g. cancer and conditions
 CC involving premature ageing). The antisense oligonucleotides of the
 CC invention are also useful for diagnostics, therapeutics and prophylaxis
 CC (e.g. to prevent or delay infection, inflammation or tumour formation).
 CC The present DNA sequence was used in the invention

XX Sequence 35000 BP; 10326 A; 6524 C; 6774 G; 11376 T; 0 U; 0 Other;

Query Match 9.4%; Score 37; DB 6; Length 35000;
 Best Local Similarity 49.2%; Pred. No. 2.6;
 Matches 97; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 192 ATAACGGGAACAGACTCATGTTGACCTGGTTGTAAGAGAGAGAGCGGTTATTATT 251
 Db 3747 ATTAGTGTCAAAAGAACCATGTTCTTATTGAAAGTCACTGCAATATATATTAC 3806
 QY 252 TCAGCATCTGTCGCCGACAGAGAGGCGCATGGAAGCCGGCGGAGACACATTCGTGTAG 311
 Db 3807 TTTCTAACATTTCCCGTGTCTAGGAATGACAGATTACTGTTGAGTGACATCATTAATAG 3866
 QY 312 ATTGATATTTAATATATAGCGTAAGTCTGCTGTTATCTATATATTATGTCATCTAAAT 371
 Db 3867 TACTATTTTAAANAATTTGTTGCATCTGTTGCTTTTATTTTAGAAGAAATTTGGAT 3926
 QY 372 CACTTTTAAAGTCAGACT 388
 Db 3927 TATTTAGAAACAAAT 3943

RESULT 8

ABL34071

ID ABL34071 standard; DNA; 13503 BP.

XX ABL34071;
 AC ABL34071;

XX 26-MAR-2002 (first entry)


```

XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 7409; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
XX Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 9.3%; Score 36.6; DB 4; Length 397;
XX Best Local Similarity 47.2%; Pred.No. 0.49;
XX Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
XX
QY 156 AAAAAATATCGTGGGAGGACAGCCCCAACCCACGATACCGGAAACAGACTCATGTG 215
DB 135 AAAAAATTATGGTTAAAGAAATAATGAAAGATCTGGCTCTTGGTCAACAGGCAAAAGACA 194
QY 216 ACCTTGGTTGTAAGAGAGAGAGCGAGCGTTATTTTTCAGCATCTGTCGCCGAGAGAG 275
DB 195 AACCTTCATCTAAAGTGATATAAATAGTTATTATACATACATCATAGCTCCAGCCTTGATC 254
QY 276 GGCAATGGAAAGCCGGCGGAGAGCAACATTGCTGTAGATTGATATTTAATATATTAGCGTA 335
DB 255 TGTGTATAGNCTGTGTTGGTCTTCATCCGAGAGGGTGGATACCCCAAGTAGAATA 314
QY 336 ACTGTTATGCTGTGTTATCTATATTATGTGATCTAAATCACTTTTAAGTCAGAGTGA 390
DB 315 TAGAGAATGACACCTGGAGGAGAATAAATGTTTAACACAATTTAAGTCATATAGA 369
XX
XX RESULT 14
XX AAK07110
XX ID AAK07110 standard; DNA; 397 BP.
XX
XX AAK07110;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7101.
XX
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.

```

```
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483446/52.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX PS Example 4; SEQ ID NO 7101; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention
XX CC
XX CC Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX CC
Query Match 9.3%; Score 36.6; DB 4; Length 397;
Best Local Similarity 47.2%; Pred. No. 0.49;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 156 AAAAATATCGTGGGAGAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTG 215
DB 135 AAAAATATCGTGGTAAAGAAATGAAAGATCTGGCTTGGTCAACAGGCAAAAGACA 194
QY 216 ACCTTGGTTGTAAGAGAGAGAGCGGCTTATTATTTCAGCATCTGTCGCGCAGAGAG 275
DB 195 AACCTTCATCTAAAGTGATATAAATAGTTTATACATACATAGCTCCAGCCTTGATC 254
QY 276 GGCATGGAAAGCCGGCGAGAGCAACATTCGCTGTAGATTGATATTTAATATTAGCGTA 335
DB 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAAAGGTGGATACCCCAAGTAGAATA 314
QY 336 ACTGTTATCTGTTATCTATATATGATCTAAATCACTTTTAAAGTCAGAGTGA 390
DB 315 TAGAGATGACACCTGGAGGAGATATAATGTTAACAATTTAAGTCATATAGA 369
RESULT 15
ABS32578
ID ABS32578 standard; DNA; 397 BP.
XX AC ABS32578;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver single exon probe, SEQ ID NO 7568.
XX KW Human: single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX
```

```
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488898/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 1; SEQ ID NO 7568; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABS25011-ABS51005 represent human
XX CC liver single exon nucleic acid probes of the invention. Note: The
XX CC sequence information for this patent does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX CC
XX CC Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX CC
Query Match 9.3%; Score 36.6; DB 4; Length 397;
Best Local Similarity 47.2%; Pred. No. 0.49;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 156 AAAAATATCGTGGGAGAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTG 215
DB 135 AAAAATATCGTGGTAAAGAAATGAAAGATCTGGCTTGGTCAACAGGCAAAAGACA 194
QY 216 ACCTTGGTTGTAAGAGAGAGAGCGGCTTATTATTTCAGCATCTGTCGCGCAGAGAG 275
DB 195 AACCTTCATCTAAAGTGATATAAATAGTTTATACATACATAGCTCCAGCCTTGATC 254
QY 276 GGCATGGAAAGCCGGCGAGAGCAACATTCGCTGTAGATTGATATTTAATATTAGCGTA 335
DB 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAAAGGTGGATACCCCAAGTAGAATA 314
QY 336 ACTGTTATCTGTTATCTATATATGATCTAAATCACTTTTAAAGTCAGAGTGA 390
DB 315 TAGAGATGACACCTGGAGGAGATATAATGTTAACAATTTAAGTCATATAGA 369
Search completed: December 18, 2004, 23:21:43
Job time : 261.841 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 15:34:11 ; Search time 43.9637 Seconds
(without alignments)
6353.876 Million cell updates/sec

Title: US-10-759-889-3
Perfect score: 393
Sequence: 1 tccaacctcaaacgacagt.....ctttaagtcagagtgaata 393
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	146.6	37.3	258	US-09-489-039A-6467	Sequence 6467, Ap
C 2	44.2	11.2	7218	US-08-232-463-14	Sequence 14, Appl
C 3	37.2	9.5	1035	US-09-489-039A-1178	Sequence 1178, Ap
C 4	34	8.7	1230025	US-09-198-452A-1	Sequence 1, Appli
C 5	33	8.4	8876	US-08-961-527-170	Sequence 170, App
C 6	32	8.1	240	US-09-252-991A-3764	Sequence 3764, Ap
C 7	32	8.1	1383	US-09-252-991A-3609	Sequence 3609, Ap
C 8	31.6	8.0	2473	US-09-173-914-3	Sequence 3, Appli
C 9	31.4	8.0	3768	US-09-248-796A-55	Sequence 55, Appl
C 10	31	7.9	2289	US-09-614-221A-375	Sequence 375, App
C 11	30.8	7.8	369	US-09-270-767-1375	Sequence 1375, Ap
C 12	30.8	7.8	369	US-09-270-767-17257	Sequence 17257, A
C 13	30.6	7.8	336	US-09-060-756-250	Sequence 250, App
C 14	30.6	7.8	336	US-09-670-314-250	Sequence 250, App
C 15	30.6	7.8	375	US-09-621-976-8473	Sequence 8473, Ap
C 16	30.6	7.8	118067	US-09-497-855A-32	Sequence 32, Appl
C 17	30.6	7.7	852	US-09-134-000C-3192	Sequence 3192, Ap
C 18	30.4	7.7	403	US-09-252-991A-7045	Sequence 7045, Ap
C 19	30.2	7.7	450	US-09-513-995C-31966	Sequence 31966, A
C 20	30.2	7.7	1284	US-09-107-532A-2979	Sequence 2979, Ap
C 21	30.2	7.7	736	US-09-270-767-7757	Sequence 7757, Ap
C 22	30	7.6	736	US-09-270-767-23039	Sequence 23039, A
C 23	30	7.6	1001	US-09-671-317-364	Sequence 364, App
C 24	30	7.6	1001	US-09-671-317-365	Sequence 365, App
C 25	30	7.6	1001	US-09-671-317-366	Sequence 366, App
C 26	30	7.6	1001	US-09-671-317-367	Sequence 367, App
C 27	30	7.6	1001	US-09-671-317-367	Sequence 367, App

C 28	30	7.6	168575	4	US-09-426-230-1	Sequence 1, Appli
C 29	29.8	7.6	3026	4	US-09-963-137-136	Sequence 136, App
C 30	29.8	7.6	3026	4	US-09-963-137-140	Sequence 140, App
C 31	29.8	7.6	3172	4	US-09-976-594-940	Sequence 940, App
C 32	29.8	7.6	6423	4	US-09-774-528-244	Sequence 244, App
C 33	29.8	7.6	33769	4	US-09-544-398B-8	Sequence 8, Appli
C 34	29.8	7.6	33769	4	US-09-543-771-8	Sequence 8, Appli
C 35	29.6	7.5	193303	4	US-09-497-855A-37	Sequence 37, Appl
C 36	29.6	7.5	193303	4	US-09-497-855A-44	Sequence 44, Appl
C 37	29.4	7.5	427	4	US-09-621-976-17978	Sequence 17978, A
C 38	29.4	7.5	474	4	US-09-621-976-18033	Sequence 18033, A
C 39	29.4	7.5	581	4	US-09-270-767-10440	Sequence 10440, A
C 40	29.4	7.5	613	2	US-08-740-309-1	Sequence 1, Appli
C 41	29.4	7.5	1200	4	US-09-252-991A-7018	Sequence 7018, Ap
C 42	29.4	7.5	1568	4	US-09-799-451-653	Sequence 653, App
C 43	29.4	7.5	1587	4	US-09-489-039A-6670	Sequence 6670, Ap
C 44	29.4	7.5	2469	4	US-08-489-039A-6749	Sequence 6749, Ap
C 45	29.4	7.5	22846	2	US-08-469-461-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-489-039A-6467/c
; Sequence 6467, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6467
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6467

Query Match	37.3%	Score	146.6	DB	4	Length	258
Best Local Similarity	78.2%	Pred. No.	4.4e-39				
Matches	176	Conservative	0	Mismatches	49	Indels	0
Gaps	0						
Qy	8	TCAAACGAACAGTCGGGATATCAAAATAAAACAAAGCAGCAATAGAGCGCGTGTGAACAA	67				
Db	258	TCAAACGAACAGTCGGGATATCAAAATAAAACAAAGCAGCAATAGAGCGCGTGTGAACAG	199				
Qy	68	CGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAAGTACCGCGCCAGCCAGCCA	127				
Db	198	TGCTGGGTGCCAGACAAATCGTAAATGCGGTCCGAACCCAGCAGTCGTCACCAAGCCA	139				
Qy	128	GAAAAATGCCAGTGATAAAAGCAATTCAAAAATATCGTGGGAAGGACAGCCCAACAC	187				
Db	138	AAAGATAGCCAGCGCAAAAGCAACTCGAAGAAGATCGCGGGAACGATTAAGCCGAAAT	79				
Qy	188	CACGATAACGGAAACAGACTCATGTGACCTTGTGTTAAAGAG	232				
Db	78	CACGATCACAGGAACAGACTCATGTGACCTTGTGTTAAAGAG	34				

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMWU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9PT-Fls
US-08-232-463-14

Query Match 11.2%; Score 44.2; DB 1; Length 7218;
Best Local Similarity 1.3%; Pred. No. 0.00035;
Matches 4; Conservative 180; Mismatches 113; Indels 0; Gaps 0;
QY 10 AACGACAGTCGCGATATCAATAAAACAGCAGCAATAGACGCGGTGTGAACAACG 69
DB 1358 RRR 1299
QY 70 CCGATGCCAGCAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACCCAGA 129
DB 1298 RRR 1239
QY 130 AATCCCGAGTATAAAGCAATCAAAAATATCGTGGGAGGACGACCCCAACCA 189
DB 1238 RRR 1179
QY 190 CGATAACGGGAAACAGACTCATGTCACCTTGTGTAAGAGAGAGACGCGTTATTAT 249
DB 1178 RRR 1119
QY 250 TTTCAGCATCTGCGCGGAGAGAGCGGATGGAACCGCGGAGAGCAACATTGC 306
DB 1118 RRR 1062

RESULT 3
US-09-489-039A-1178/c
; Sequence 1178, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 1178
LENGTH: 1035
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1178
Query Match 9.5%; Score 37.2; DB 4; Length 1035;
Best Local Similarity 50.0%; Pred. No. 0.026;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 63 AACACGCCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACCC 122
DB 731 AACGCTGGCGCCGCCGCAACATAGTCCTAGTTCCGGGTACGGAGAAATTCGCGCGCACCC 672
QY 123 AGCAGAAATCCCAAGTGTATAAAGCAATTCAAAAAATATCGTGGGAGGACAGCCCA 182
DB 671 ACGCCACCCAGCGTCATCCAGCCAAAGAGAACGGTGATTCGACGACGCCACAGAGCCCC 612
QY 183 AACACACCATATAACGGGAAACAGACTCATGTTGACCTTGTGTAAGAGAGAGAGCGGCG 242
DB 611 GGCTGCACGACGCTGGAGAGCAGGATAATCAGGAACAGCGTTGGCATGCCGACCACT 552
QY 243 TTATTA 248
DB 551 TCGATA 546

RESULT 4
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature


```

; NAME/KEY: misc_feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature

Query Match      8.7%; Score 34; DB 4; Length 1230025;
Best Local Similarity 52.9%; Pred. No. 14; Indels 0; Gaps 0;
Matches 73; Conservative 0; Mismatches 65;

QY      79 AGACAAAGTCGTAGATACCTGTTGGACAAAGTACCCGGCGCACACCACCCAGAAATCGCCA 138
Db      1176718 AGAGAAATATCTGTAAGCAATTTGGCAAAGATAGCGCGTTAAACAGCAAGCAAAATTCGA 1176659

QY      139 GTGATAAAGCAATTCAAAATATATCGTGGAGACGACCCCAACACACACGATACCG 198
Db      1176658 ATAAAGAAAGTAAATCGAAGAGATATCGATTTTAGTCAAAATATGAGAACTCGCAATGG 1176599

QY      199 GAAACAGACTCATGTTGA 216
Db      1176598 TAACGATCTCTACCATGA 1176581

```

RESULT 5
 US-08-961-527-170/c
 ; Sequence 170, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PS340P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 170:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8876 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double

; TOPOLOGY: linear
US-08-961-527-170

Query Match 8.4%; Score 33; DB 4; Length 8876;
Best Local Similarity 51.9%; Pred. No. 2.1;
Matches 98; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 179 CCCAAACACCAGGATAACGGGAACAGACTCATGTGTGACCTTGGTTGTAACAGAGAGCA 238
DB 4568 CCCAACGATTCGCGGAGAAAATGATAAACCCTGCAGGTGTGATATAAACTTCAATCA 4509
QY 239 GCGCTTATTATTTTCAGCMTCTGTCGCCGACAGAAGGCGCATGGAAGCCGGCGAGAGC 298
DB 4508 GCCTTGACACAATGCTTCTCATCAAGACGAAGGGAAGGCG-TTCAATGATAATGGAGAGA 4450
QY 299 AACATTGCTGTAGATGATATTTATATATATTAGCGTAACTGTTATGCTGTTATCTATATT 358
DB 4449 AAAATGGCTCCAAGTTGAAGTACACTAGATGGAGTGTGTTAAAAAGCAGTCATCTAGTTT 4390
QY 359 ATGTGATCT 367
DB 4389 TTTCTTTCT 4381

RESULT 6
US-09-252-991A-3764/c
; Sequence 3764, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3764
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3764

Query Match 8.1%; Score 32; DB 4; Length 240;
Best Local Similarity 49.4%; Pred. No. 0.64;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 37 ACAAGCAGCAATAGAGCGCGGTGTGAAACAAGCCGGATGCCAGACAAAGTCGTAGATAC 96
DB 208 ACAGGCAGCGCAACAGCGCTGACCCCGAACAGCGCCGGGTGCCAGGTGGAACGGATACAGGC 149
QY 97 CTGTTGGCACAGTAGTCCC GGCGCACCCAGCGAGAAAATCGCCAGTGTATAAAGCAATTCAA 156
DB 148 CAAACGAGCGGAGATCTCCGTGAGCCCCCAGGTGAGCCCGCGGCGATCACGAACAACA 89
QY 157 AAAAATATCGTGGGAGGACAGCCCAAAACACCACGATACGGGAAACA 204
DB 88 GCGTCAGGCTCGCATGTAGACCCCATGGATGGCGATTTTACAGGAGCA 41

RESULT 7
US-09-252-991A-3609
; Sequence 3609, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3609
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3609

Query Match
Best Local Similarity 8.1%; Score 32; DB 4; Length 1383;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 37 ACAAGCAGCAATAGAGCGCGGTGTTGAACAAGCCCGGATGCCAGACAAGTCTAGATAC 96
Db 1137 ACAGGCGGGAACAGGCTGACCCGGAACAGCGCGGTCGCCAGTGAGCGATACAGGC 1196

QY 97 CTGTTGGCAAGTACCCCGGCGCACCAGCCAGCAAAATCGCAGTGTAAAGCAATTCAA 156
Db 1197 CAACCGAGCGAAGATCCGGTCGAGCCCGCAGGTGAGCGCCGCGGATCAGAACAA 1256

QY 157 AAAATATCGGTGGGAGGACAGCCCAACACACCAATACGCGGAACA 204
Db 1257 GCGTCAGGTCGGCATGTAGACCCCATGGATCGCGATTTCACGAGCA 1304

RESULT 8
US-09-173-914-3
; Sequence 3, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857el Leucine zipper, KARP-1 and
; FILE REFERENCE: B0877/7017/HK
; CURRENT APPLICATION NUMBER: US/09/173,914
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 60/064,557
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Mesocricetus Auratus
US-09-173-914-3

Query Match
Best Local Similarity 8.0%; Score 31.6; DB 3; Length 2473;
Matches 79; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 155 AAAAATATCGTGGGAGGACAGCCCAACACACAGATACGGGAACAGACTCATGTT 214
Db 1673 AAAAAGGATCAAGTACGCTCAGGACATTTTCCAAGACAATGATGAGAGGGCGCTGCT 1732

QY 215 GACCTTGTGTTAAAGAGAGACAGCGCGTATTATTTTCAGCATCTGTCGCGGAGAGAA 274
Db 1733 GCCAAAAGTGTAAATGGAGAAAGAGAGTACATCATCTCTAGCCTGGCTGAA 1792

QY 275 GGGCATGGAAGCCGGCGGAGAGCAACATTTGCTGTAGA 312
Db 1793 GGGAAATGTCAACAGGTTTGAAGTGTGAATCTCTGTGA 1830

RESULT 9
US-09-248-796A-55/C
; Sequence 55, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 55
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-55

Query Match
Best Local Similarity 8.0%; Score 31.4; DB 4; Length 3768;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 302 ATTGCTGTAGATTGATTTTAAATATATATAGCGTAACCTGTTATCTTATCTATATG 361
Db 3137 AATGATCCAAATTTGATATTATTTTCATATAATTTAAATCACTTAAATTTAATATA 3078

QY 362 TGATCTAATCACTTTTAACTCAGAGTGA 390
Db 3077 GGATCTAAATCTGGTGTCTCTGATTGTA 3049

RESULT 10
US-09-614-221A-375/C
; Sequence 375, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 375
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-375

Query Match
Best Local Similarity 7.9%; Score 31; DB 4; Length 2289;
Matches 76; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 134 CGCAGTGTATAAAGCAATTCAAAAATATCGTGGGAAGGACAGCCCAACACACCGAT 193
Db 2211 CACTAGTGGTGAAGATATACAAAAACCCACGCAACAGCAGGACACACACAGAA 2152

QY 194 AACGGAAACAGACTCATGTTGACCTTGTGTTGTTAAAGAGAGAGCGGCTTATTTTC 253
Db 2151 GACGGTATAGGGTTTAACTAGACCTTAGGTTCCGTTTCATATGATGCACCTGAAATATC 2092

QY 254 AGCATCTGTCGCGCAGAGAGGCGCATGAA 284
Db 2091 CTGCTTTTCTCCAAATCCATGATTTGCA 2061

RESULT 11
US-09-270-767-1975
; Sequence 1975, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
```


Search completed: December 18, 2004, 20:49:00
Job time : 48.9637 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:42:54 ; Search time 272.694 Seconds
(without alignments)
7956.501 Million cell updates/sec

Title: US-10-759-889-3
Perfect score: 393
Sequence: 1 ttcaacctcaaacgacagct.....cttttaagtcagagtgaata 393

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4093002 seqs, 2760418825 residues

Total number of hits satisfying chosen parameters: 8186004

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	393	17	US-10-759-889-3
2	39	9.9	1842	18	US-10-425-115-137044
3	38.4	9.8	1598	17	US-10-437-963-70841
4	37	9.4	35000	10	US-09-793-807-10
5	36.8	9.4	13503	15	US-10-311-455-2044
6	36.6	9.3	397	9	US-09-864-761-6274
7	36.6	9.3	330973	13	US-10-087-192-1498
8	35.6	9.1	711	16	US-10-335-977-4100
9	35.6	9.1	714	16	US-10-335-977-4099
10	35.6	9.1	723	16	US-10-335-977-4101
11	35	8.9	513509	10	US-09-754-853A-4
12	34.4	8.8	83391	17	US-10-433-793-124

ALIGNMENTS

RESULT 1

US-10-759-889-3
; Sequence 3, Application US/10759889
; Publication No. US20040157331A1
; GENERAL INFORMATION:
; APPLICANT: Van Dyk, Tina K
; TITLE OF INVENTION: Regulator/Promoter for Tunable Gene Expression and Metabolite
; TITLE OF INVENTION: Sensing
; FILE REFERENCE: CL2036 US NA
; CURRENT APPLICATION NUMBER: US/10/759,889
; CURRENT FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: 60/440,965
; PRIOR FILING DATE: 2003-01-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-759-889-3

Query Match 100.0%; Score 393; DB 17; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.7e-110;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTCAACCTCAACGACGACGATCAATAAACAAGCAGCAATAGAGCGCGTGT 60
Db 1 TTCAACCTCAACGACGACGATCAATAAACAAGCAGCAATAGAGCGCGTGT 60

Qy 61 TGAACAACGCGGATCCGACAAAGTCGTAGATACCTGTGGCAACAAGTACCCGCGCA 120
Db 61 TGAACAACGCGGATCCGACAAAGTCGTAGATACCTGTGGCAACAAGTACCCGCGCA 120

Qy 121 CCAGCCAGAAATCGCCAGTGAATAAGCAATTCAAAAATATCGGTGGGAAGGACGCC 180

Sequence 262484,
Sequence 262484,
Sequence 1, Appli
Sequence 214898,
Sequence 214898,
Sequence 2, Appli
Sequence 99, Appli
Sequence 32, Appli
Sequence 52886, A
Sequence 52886, A
Sequence 23619, A
Sequence 163667, A
Sequence 18587, A
Sequence 117791, A
Sequence 26057, A
Sequence 26058, A
Sequence 182312,
Sequence 182312,
Sequence 8255, Ap
Sequence 8607, Ap
Sequence 37362, A
Sequence 170, App
Sequence 170, App
Sequence 287219,
Sequence 287219,
Sequence 24681, A
Sequence 129690,
Sequence 6870, Ap
Sequence 2926, Ap
Sequence 41098, A
Sequence 682, App
Sequence 1034, Ap
Sequence 33477, A

Db 121 CCAGCCAGAAATCCCGAGTGTATTAAGCAATTCACAAATATCGGTGGGAAGACAGCC 180
QY 181 CAAACACACAGTAACGGGAACAGACTCATGTTGACCTTGGTCTTAAGAGAGAGCAGG 240
Db 181 CAAACACACAGTAACGGGAACAGACTCATGTTGACCTTGGTCTTAAGAGAGAGCAGG 240
QY 241 CGTTATTTATTTTTCAGCATCTGTGCGCGGAGAGAGGCGATCGAAAGCGCGGCGAGAGCAA 300
Db 241 CGTTATTTATTTTTCAGCATCTGTGCGCGGAGAGAGGCGATCGAAAGCGCGGCGAGAGCAA 300
QY 301 CATTGCTGTAGATTGATATTTAATATATATAGGTAACGTTATGCTGTATCTATATAT 360
Db 301 CATTGCTGTAGATTGATATTTAATATATATAGGTAACGTTATGCTGTATCTATATAT 360
QY 361 GTGATCTAATCACTTTTAAAGTCAGAGTGAATA 393
Db 361 GTGATCTAATCACTTTTAAAGTCAGAGTGAATA 393

RESULT 2

US-10-425-115-137044
; Sequence 137044, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 137044
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MR74577_56460C.1
US-10-425-115-137044

Query Match 9.9%; Score 39; DB 18; Length 1842;
Best Local Similarity 58.0%; Pred. No. 0.41; Mismatches 50; Indels 0; Gaps 0;
Matches 69; Conservative 0;

QY 67 ACGCCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAGTACCCGGCGCACAGCC 126
Db 1574 ACAACAGTACCCAACTAGCTCAAACTTCCTTGTGGCCTAAACTAAGAGATCCTCCAGCC 1633
QY 127 AGAAATCCCGAGTGTATTAAGCAATTCACAAATATCGGTGGGAAGCAGACGCCAAAC 185
Db 1634 CTTAGATCCCGAGGAGATCAACACAAATATCTGTACCAAGGAGCGTCCATGAC 1692

RESULT 3

US-10-437-963-70841
; Sequence 70841, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 70841
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71373C.1
US-10-437-963-70841

Query Match 9.8%; Score 38.4; DB 17; Length 1598;
Best Local Similarity 56.2%; Pred. No. 0.58; Mismatches 56; Indels 0; Gaps 0;
Matches 72; Conservative 0;

QY 126 CAGAAATCCCGAGTGTATTAAGCAATTCACAAATATCGGTGGGAAGAGCAGCCAAAC 185
Db 1161 CACAGAACACTAATAATAACAGCAACGGGAACATATGTAGGTGGCAGCCACAAACAGAG 1220
QY 186 ACCACGATACCGGAACAGACTCATGTTGACCTTGTAAAGAGAGAGCAGCGTTA 245
Db 1221 GCTAACTCTAGTGAGAGAAATCATGAGAAATACGACAAAGATGAGATGAGGTGTTTC 1280
QY 246 TTATTTTC 253
Db 1281 TTCTTATC 1288

RESULT 4

US-09-793-807-10
; Sequence 10, Application US/09793807
; Publication No. US20030171310A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL EXPRESSION
; FILE REFERENCE: RTS-0206
; CURRENT APPLICATION NUMBER: US/09/793,807
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 35000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-793-807-10

Query Match 9.4%; Score 37; DB 10; Length 35000;
Best Local Similarity 49.2%; Pred. No. 6.8; Mismatches 100; Indels 0; Gaps 0;
Matches 97; Conservative 0;

QY 192 ATAACGGGAACAGACTCATGTTGACCTTGGTGTAAAGAGAGAGCAGCGTTATTTT 251
Db 3747 ATTAGTGTCAAAAGAACCAATGTTGTTCTTTATTGAAAGGTCACTGCAATATATAATTTAC 3806
QY 252 TCAGCATCTGTCGCCAGAGAGGCGCATGGAAGCCGGCGGAGAGACATTTGCTGTAG 311
Db 3807 TTTCTAACATTTTCCGTGCTTAGGAATGACAGATTACTGTTGAGTGAACATCAATTAATAG 3866
QY 312 ATTGATATTTAATATATATTAGCGTAACTGTTATGCTGTTATCTATATATTGATCTAAAT 371
Db 3867 TACTATTTTAAAACTTGTGTCACCTGTTGCTTTTATTTTAGAAGAAATTTGGAT 3926
QY 372 CACTTTTAAGTCAGACT 388
Db 3927 TATTTTAGAACAAAT 3943

RESULT 5

US-10-311-455-2044
; Sequence 2044, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander


```
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2044
; LENGTH: 13503
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2044

Query Match          9.4%; Score 36.8; DB 15; Length 13503;
Best Local Similarity 63.6%; Pred. No. 5;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 299 AACATTGCTGTAGATGATTTATATATATTAGCGTAACCTGTATGCTGTTATCTATATT 358
Db 10532 AATATTGATGAAGTAGATTTTATTTAATAGTGTAAATTTTAGTTTTTTTAAATT 10591

QY 359 ATGTGATCTAAATCACTTTTAAGTCAGA 386
Db 10592 TAGTTAATAAATTAGTGTAAATTCGA 10619

RESULT 6
US-09-864-761-6274
; Sequence 6274, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6274
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004056.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
US-09-864-761-6274
```

```
Query Match          9.3%; Score 36.6; DB 9; Length 397;
Best Local Similarity 47.2%; Pred. No. 1.1;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 156 AAAAAATATCGTGGGAAGGACGCCAAACACCCACGATAACGGGAACACAGACTCATGTTG 215
Db 135 AAAAAATTATGTTTAAAGAAAAAATGAAAGATCTGGCTCTTGTGTCACAGCAAAAGACA 194

QY 216 ACCTTGGTGTAAAGAGAGAGAGCGCGCTTATTATTTTCAGCATCTGTCGGCGGAGAGAAG 275
Db 195 AACCTTCATCTAAAGTGATATAAATAGTTTATTCATACATCAATAAGTCCAGCCTTGTATC 254

QY 276 GGCATGGAAGCCGGGCGAGAGCAACATTGCTGTAGATTGATATTTAATATATTAGCGTA 335
Db 255 TGCTGATAGAGTCTGTGTGTTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 314

QY 336 ACTGTTATGCTGTATCTTATATATATATGATGATCAATCACTTTTAAGTCAGAGTGA 390
Db 315 TAGAGAAATGACACCTGGAGGAGATATAATGTTTAACAACAATTTAAGTCATATAGA 369
```

```
RESULT 7
US-10-087-192-1498/c
; Sequence 1498, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1498
; LENGTH: 330973
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(330973)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1498

Query Match          9.3%; Score 36.6; DB 13; Length 330973;
Best Local Similarity 47.2%; Pred. No. 26;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 156 AAAAATATCGTGGGAGGACGACCCCAACACACACGATACCGGAAACAGAGCTCATGTTG 215
Db 254607 AAAAATATGTTTAAAGAAATGAAAGATCTGGCTCTTGGTCAACAGGCAAGAGACA 254548

QY 216 ACCTTGGTTGTAAGAGAGAGCAGGCGTTATTATTTTCAGCATCTGTCCGCCAGAGAG 275
Db 254547 AACCTTCATCAAGTGATATAAATAGTTATTACATACATCATAGCTCCAGCCTTGATC 254488

QY 276 GGCATGGAAGCGCGGAGAGCAACATTGCTGTAGATTGATTTTATATATTAGCGTA 335
Db 254487 TGTGTATAGAGTCTGTGTGGTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 254428

QY 336 ACTGTTATGCTGTATCTATATATATGATGATCTAAATCACTTTTAAAGTCAGAGTGA 390
Db 254427 TAGAGATGACACCTGGAGAGATATATATGTTAAACACAATTTAAGTCATATAGA 254373

RESULT 8
US-10-335-977-4100
; Sequence 4100, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4100:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori

;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...714
; SEQUENCE DESCRIPTION: SEQ ID NO: 4100:
US-10-335-977-4100
```

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...711
; SEQUENCE DESCRIPTION: SEQ ID NO: 4100:
US-10-335-977-4100

Query Match          9.1%; Score 35.6; DB 16; Length 711;
Best Local Similarity 51.9%; Pred. No. 2.9;
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 30 AAATAAACAAGCAGCAATAGAGCGGGTGTTCGAACAACCGCCGATGCCAGACAAAGTCG 89
Db 344 AATCATCTCCTTAGACCAAAAGACTGCTGAGCAACAATGCGGAATATCGATAAGTTT 403

QY 90 TAGATACCTTTGGCACAAGTACCGCGGCACACGACCAAAAATCGCAGTGATAAAAGC 149
Db 404 TAGTTCTTTTATTGTGAGGGAGAGCTGAAAATCGAGCAAGAAAAACAAATATAGAGAGC 463

QY 150 AATTCAAAAATATATCGTGGGAAAGGACAGCCAA 183
Db 464 GTTTAGAAATTTCTTAAGGCATATCAGAGCAA 497

RESULT 9
US-10-335-977-4099
; Sequence 4099, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4099:
SEQUENCE CHARACTERISTICS:
LENGTH: 714 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori

;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...714
; SEQUENCE DESCRIPTION: SEQ ID NO: 4099:
US-10-335-977-4099
```

US-10-335-977-4099

Query Match 9.1%; Score 35.6; DB 16; Length 714;
Best Local Similarity 51.9%; Pred. No. 2.9;
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 30 AATAAACAAGCAGCAATAGACCGCGGTGTTGAACAACCGCGGATGCCAGACAAAGTCG 89
DB 344 AATCAATCCCTTAGACCAAAAGACTGCTGAGCAACAATCGCGAATATCCGATAAAGTTT 403
QY 90 TAGATACCTGTTGGCACAAGTACCGCGCCACACAGCCAGCAAAATCCCGAGTATAAAGC 149
DB 404 TAGTTCTTTATTGTGAGGGAAGCTGAAATCGAGCAAGAAACAAATATAAGAGAGC 463
QY 150 AATCAAAAATATCGTGGGAGGACAGCCCAA 183
DB 464 GTTAGAAACTCTCTAAAGGCATATCAGAGCAA 497

RESULT 10

US-10-335-977-4101
; Sequence 4101, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0

SOFTWARE: UNIX

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 4101:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...723

SEQUENCE DESCRIPTION: SEQ ID NO: 4101:
US-10-335-977-4101

Query Match 9.1%; Score 35.6; DB 16; Length 723;
Best Local Similarity 51.9%; Pred. No. 2.9;

Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 30 AATTAACAAGCAGCAATAGACCGCGGTGTTGAACAACCGCGGATGCCAGACAAAGTCG 89
DB 356 AATCAATCCCTTAGACCAAAAGACTGCTGAGCAACAATCGCGAATATCCGATAAAGTTT 415
QY 90 TAGATACCTGTTGGCACAAGTACCGCGCCACACAGCCAGCAAAATCCCGAGTATAAAGC 149
DB 416 TAGTTCTTTATTGTGAGGGAAGCTGAAATCGAGCAAGAAACAAATATAAGAGAGC 475
QY 150 AATCAAAAATATCGTGGGAGGACAGCCCAA 183
DB 476 GTTAGAAACTCTCTAAAGGCATATCAGAGCAA 509

RESULT 11

US-09-754-853A-4/c
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li

TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max

NAME/KEY: CDS
LOCATION: (111805)..(113968), (114684)..(115204)

FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(513509)

OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 318013_region_A3

US-09-754-853A-4

Query Match 8.9%; Score 35; DB 10; Length 513509;
Best Local Similarity 47.9%; Pred. No. 1e+02;
Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 182 AATCAACGATTAACCGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGAGCGC 241
DB 275036 AATAACAAATCAACCCGAAATATAACAAATTAAGCTTAAGCTTAAATCAAAAAA 274977

QY 242 GTTATTATTTTCAGCATCTCTCGCCGAGAGAGAGGCGATCGAAAGCGCGGAGAGCAAC 301
DB 274976 ATATTATTATTAATTTTATAAGCAAAAAAGGCAATGATATATTGATAAAATGAT 274917

QY 302 ATTGCTGTAGATTGATTAATTAATATTATAGCGTAACGTGTTATCTGTTATCTATTATG 361
DB 274916 ATATATGTATATAGATCAATAATTTTGAACAAAAATCTAAACATGTAATCTTTTAT 274857

QY 362 TGATCTAAATCACTTTTAAAGTCAGAGTGAAT 392
DB 274856 TTCTTTAAATTCACITTAGGCTGTTGGAT 274826

RESULT 12

US-10-433-793-124/c

; Sequence 124, Application US/10433793

; Publication No. US20040142334A1

; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG


```
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
```

Query Match 8.7%; Score 34; DB 16; Length 1230025;

Best Local Similarity 52.9%; Pred. No. 3e+02;

Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
QY 79 AGCAAAAGTCGTAGATACCTGTTGGCACAAGTACCGCGGCGGACAGCCAGCCAGAGAAATCGCCA 138
Db 1176718 AGCAAAATCTCTAAGCAAAATTCGCAAAAGATAGCGCGTTAACAGCAAGCAAAATTCOA 1176659

QY 139 GTGATAAAGCAATTCAAAATAATCGTGGGAAGGACAGCCCAACACCAACGATAACGG 198
Db 1176658 ATAAGAACGTAAATCGAAGATATCGATTTTAGTCAAAATATGAGGAATCGCAATGG 1176599

QY 199 GAAACAGACTCATGTTGA 216
Db 1176598 TAACGATCCTCACCATGA 1176581
```

Search completed: December 18, 2004, 23:07:19
Job time : 279.694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 22:46:01 ; Search time 1710.72 Seconds
(without alignments)
8371.204 Million cell updates/sec

Title: US-10-759-889-3

Perfect score: 393

Sequence: 1 ttcaacctcaaacgaacagt.....ctttaagtcagagtgaata 393

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	288.2	73.3	755	8	AF094919
2	162.8	41.4	785	9	CL663057
3	40.2	10.2	942	9	CN80433N
4	39.6	10.1	1101	9	CN80021D
5	39	9.9	790	6	CB665130
6	38.4	9.8	424	5	BU673561
7	38.4	9.8	453	6	CF324862
8	38.4	9.8	727	2	BB657476
9	38.4	9.8	846	9	CN8033M4
10	38	9.7	707	6	CD849459
11	38	9.7	755	8	BH485998
12	37.8	9.6	417	1	AL389174
13	37.8	9.6	524	9	CE225599
14	37.8	9.6	3190	3	AK085573
15	37.2	9.5	603	8	AZ230411
16	37.2	9.5	665	4	BG546513
17	37.2	9.5	997	9	CN8005TE
18	36.8	9.4	922	9	CN80073W
19	36.6	9.3	296	7	CO253598
20	36.6	9.3	812	9	CC552656
21	36.6	9.3	822	8	BZ603536
22	36.4	9.3	616	9	CE401841
23	36.4	9.3	861	8	AZ528603
24	36.2	9.2	406	1	AL389175

25	36.2	9.2	593	5	BO702511
26	36.2	9.2	840	9	AG532072
27	36.2	9.2	1000	7	CR278250
28	36	9.2	652	9	CB662916
29	35.8	9.1	928	9	CNS0020M
30	35.8	9.1	293	2	BB193716
31	35.8	9.1	665	9	CE011345
32	35.6	9.1	238	6	CB279545
33	35.6	9.1	675	9	CC946844
34	35.6	9.1	677	8	BZ088375
35	35.6	9.1	756	8	BZ451202
36	35.6	9.1	1139	8	CC202232
37	35.6	9.1	1302	8	BZ556775
38	35.2	9.0	679	5	BU484568
39	35.2	9.0	731	8	BZ049478
40	35.2	9.0	867	7	CR289354
41	35.2	9.0	1044	9	CL509018
42	35	8.9	895	5	BU152641
43	35	8.9	913	7	CO081762
44	35	8.9	922	7	CK016766
45	35	8.9	1003	4	BG179376

ALIGNMENTS

RESULT 1
AF094919
LOCUS AF094919 755 bp DNA linear GSS 29-AUG-2000
DEFINITION AF094919 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 906-T3, genomic survey sequence.
ACCESSION AF094919
VERSION AF094919.1 GI:4322761
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 755)
AUTHORS Wong, R.M.Y. and McClelland, M.
TITLE End Sequences of Salmonella typhimurium LT2 Lambda DASHII Clones, Li-Cor
JOURNAL Unpublished (1999)
COMMENT Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.

FEATURES

source
Location/Qualifiers
1..755
/Organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone_lib="906-T3"
/clone="906-T3"
/note="Vector: Lambda DASH II; sequenced using Li-Cor sequencer"

ORIGIN

Query Match 73.3%; Score 288.2; DB 8; Length 755;
Best Local Similarity 85.0%; Pred. No. 1.8e-75;
Matches 334; Conservative 0; Mismatches 58; Indels 1; Gaps 1;
QY 1 TTCAACCTCAACGAACAGTCGGGATATCAATAAACAAGCAGCAATAGAGCGGTGT 60
DB 189 TTCAACCTCAACGAACAGTCGGGATATTAATAAACAAGCAGCAATAGAGCGGTAT 248
QY 61 TGAACAACGCGGATGCCAGCAAAAGTCGTAGATACCTTTGGCACAAGTACCGCGGCA 120
DB 249 TGAACAAGCGGATGCCAGCAAAATCATAGATGCCCGCGGCAACGATCCGACGCA 308

[illegible]

Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

source
1. .727
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230002J12"
/issue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo
eyeball"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCAACTCGAGTGTGTTTTTTTTTTN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATCTCGATTAAATTAATATCCCGCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBlueScript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 9.8%; Score 38.4; DB 2; Length 727;
Best Local Similarity 49.5%; Pred. No. 2.6;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 32 ATAAACAAGCGCAATAGAGCGCGGTGTGTAACAACCGCGGATGCCAGACAAAGTCGTA 91
Db 522 ATACAACTAACAAAATAAGATATAGAGTAAATAAACAACAGAGTAGAGACGATATATTA 463
QY 92'GATACCTGTGGCACAAGTACCGCGGCACCGCCAGCAAAATCGCCAGTGTAAAGCAA 151
Db 462 GAAATAAATATAGATAGTAGGAGGGGAGAGAGGAATATAAATAATAAATTGAA 403
QY 152 TTCAAAAAATATCGGTGGGAAGGACAGCCCAACACCAAGTAAACGGGAAACAGACTCAT 211
Db 402 TTCAATAAGTACACGTTGAAATAAACAAGAGTTTAAAAAATACTGTAAAGTCACCTAT 343
QY 212 GTTGACCTTGGTTGTAAGA 231
Db 342 AATACCTCGTTTCATAAACA 323

RESULT 9

CNS033M4 846 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION sequence.
ACCESSION AL226309
VERSION AL226309.1 GI:7985226
KEYWORDS GSS; genome survey, sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE

AUTHORS
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Saurin, W., Bernot, A., Wincker, P., Brottier, P., Quetier, F.,
Barnot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.

TITLE

Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence

JOURNAL

MEDLINE

PUBMED

20296633

10835645

AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

JOURNAL

MEDLINE

PUBMED

20359837

10899143

AUTHORS

3 (bases 1 to 846)

Genoscope.

Direct Submission

TITLE

JOURNAL

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

COMMENT

This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

1. .846

Location/Qualifiers

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone_lib="G"

/note="Genoscope sequence ID : COAG209BB09SP1-end ;

PUC-Ori"

ORIGIN

Query Match 9.8%; Score 38.4; DB 9; Length 846;
Best Local Similarity 40.9%; Pred. No. 2.7;
Matches 81; Conservative 23; Mismatches 94; Indels 0; Gaps 0;
QY 9 CAAACGAAACAGTCGCGATATCAATATAAACAAGCAGCAATAGAGCGCGGTGTTTCAACAAC 68
Db 213 CAAARAAAAAAGAAAGAAAGAAAGCAGAGGAAACCAACCAAGAAAGAAAGAAAG 272
QY 69 GCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCGCGCGCACAGCCAG 128
Db 273 GAGGAGCAGACAMWAGAGMGCGAGMAGMAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 332
QY 129 AAAATCCGAGTCGATATAAAGCAATTCAAAAAATATCGTGGGAGGAGGAGGAGGAGGAG 188
Db 333 AAAMAGCMCAGAAACCCAGGCGACCAACCGCAGCGGAAAGAAAGAAAGAAAGAAAG 392
QY 189 ACCGATACGGGAAACAGCA 206
Db 393 AAGAAACAGACAGCAACA 410

RESULT 10

CD849459/c

LOCUS

DEFINITION

CD849459

707 bp mRNA linear EST 11-JUL-2003

CD849459

CD849459.1 GI:32533281

CD849459

CD849459.1

GI:32533281

EST.

KEYWORDS

Helianthus annuus (common sunflower)

SOURCE

Helianthus annuus

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE

AUTHORS

JOURNAL

COMMENT

1 (bases 1 to 707)
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
and <http://genoplante-info.infobiogen.fr>.

FEATURES

source

1..707
Location/Qualifiers
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="p5c8"
/db_xref="taxon:4232"
/clone="HadeVR2020D12"
/tissue_type="terminal bud"
/clone_lib="HadeVR2"

ORIGIN

Query Match 9.7%; Score 38; DB 6; Length 707;
Best Local Similarity 50.3%; Pred. No. 3.5;
Matches 92; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 112 CCGCGCGCCAGCCAGAGAAATGCCAGTGTATAAAGCAATTCAAAAATATCGGTGGGA 171
DB 564 CCGAAGTCAAGGAAACAAAGACGGCGGTTCACCATCTAGCAAACTACATCGGAGCCA 505
QY 172 AGACAGAGCCCAACACACAGTAAACGGAACAGACTCATGTGACCTGGTGTAAAGA 231
DB 504 TGGGCAACCCACGTACACACAGACAGAAAGATAGACATGCGGACCGGTGATACGA 445
QY 232 GAGAGCAGGCGTTATTATTTTCAGCATCTGTGCGCGCAGAGAGGCGCATGGAAGCCGGG 291
DB 444 AGCGGAGAGATAGACATGACGCGCGCGGTGTTACTAAGAGAGCGGAGACGGGTGA 385
QY 292 CGA 294
DB 384 CGA 382

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BH485998/c
BOHIH37F BOHI Brassica oleracea genomic clone BOHIH37, genomic
survey sequence.
BH485998
BH485998.1 GI:17694102
GSS
Brassica oleracea
Brassica oleracea

REFERENCE

AUTHORS

JOURNAL

COMMENT

1 (bases 1 to 755)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOHIH37F
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TP
Class: sheared ends.

FEATURES

source

1..755
Location/Qualifiers
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHIH37"
/clone_lib="BOHI"
/notes="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match 9.7%; Score 38; DB 8; Length 755;
Best Local Similarity 65.1%; Pred. No. 3.5;
Matches 56; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 301 CATTGCTGTAGATGATTTAATATATATAGCTAACTGTTATGCTGTATCTATATTAT 360
DB 234 CATACACGTATATTAATTAATTTAATAATAATAAGTAAATATTTATTTATTTAT 175
QY 361 GTGATCTAATCACTTTTAACTCAGA 386
DB 174 AGGTTTAAATCTTTTATTTATTA 149

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AL389174
MBC53C03F1 MtBC Medicago truncatula cDNA clone MtBC53C03 T3, mRNA
sequence.
AL389174
AL389174.1 GI:9688925
EST.

REFERENCE

AUTHORS

JOURNAL

COMMENT

Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
Journet, B.P., Crespeau, H., van Tuinen, D., Gouzy, J., Jaillon, O.,
Niebel, A., Carreau, V., Chatagnier, O., Kahn, D.,
Gianinazzi-Pearson, V. and Gamas, P.,
Medicago truncatula ESTs from endomycorrhizal roots
Unpublished (2000)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrf@genoscope.cns.fr, Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
Biologie Moléculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
Mt-est@toulouse.inra.fr Website :
<http://sequence.toulouse.inra.fr/Mtruncatula.html>).
Location/Qualifiers
1..417
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MtBC53C03"
/tissue_type="arbuscular mycorrhiza"
/dev_stage="harvested 3 weeks post inoculation with Glomus
intraradices"
/clone_lib="MtBC"

FEATURES

source

1..417
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MtBC53C03"
/tissue_type="arbuscular mycorrhiza"
/dev_stage="harvested 3 weeks post inoculation with Glomus
intraradices"
/clone_lib="MtBC"
/notes="Vector: pBluescript pSK; Site 1: EcoRI; Site 2:
XhoI; M. truncatula sterilised seeds were germinated for
72h at 25 C, before transplanting into a 1/3 Epsilones soil
: 2/3 calcined Terragreen mix in the presence of onion
root fragments colonized by the arbuscular mycorrhizal
fungus Glomus intraradices (Schenck & Smith, isolate

LPAB). The plants were watered every day and twice a week with a modified nutrient Long Ashton solution without phosphate but with a high level of nitrate. After 3 weeks RNA was extracted from whole root systems. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into Uni-zap XR vector from Stratagene and packaged using Gigapack Gold packaging extracts. Plasmids containing cDNA inserts were mass-excised from phage stocks using ExAseit helper phage and propagated in SOLR cells. Clone ordering and sequencing was performed by the Centre National de Sequencage (Genoscope, Evry, France). Note : EST may be of fungal origin."

ORIGIN

Query Match 9.6%; Score 37.8; DB 1; Length 417;
Best Local Similarity 48.0%; Pred. No. 3.6; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 117;

QY 156 AAAATATCGTGGGAGGACAGCCCAACACACGATACCGGAAACAGACTCATGTTG 215
DB 162 AAAATATCGGAGTAAAGATTACAGGCAAACTACAATAACAGGAAACTCTCATGTTT 221
QY 216 ACCTCGTGTAAAGAGAGAGAGCGGTTATATTTTCAGCATCTGTGCGCCAGAGAG 275
DB 222 ACCCTACCGTGGCACCATTACTCTTTGATGCTGATCAACACATTGACGGATTGAA 281
QY 276 GGCATGAAAGCGGCGGAGAGCAACATTCGTGTAGATTGATTTAATATATTAGCGTA 335
DB 282 GACGATGGAGTAGTTACATTAATCAATTAATTTTGTAGTTGAGATTTCTTTTCAATA 341
QY 336 ACTGTTATCGTGTATCTATATATTATGTGATCTAAATCACTTTAA 380
DB 342 TAGTTATTTTAAATTTAATAATTATCAATTAATAAACAATTGAA 386

RESULT 13
CE225599/c

LOCUS
DEFINITION
tigR-gss-dog-17000338761286 Dog Library Canis familiaris genomic,
genomic survey sequence.

ACCESSION
VERSION
CE225599.1 GI:35381314

KEYWORDS
GSS

SOURCE
Canis familiaris (dog)

ORGANISM

Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 524)
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

22875432

14512627

COMMENT

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

1..524

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood."

FEATURES

source

ORIGIN

Query Match 9.6%; Score 37.8; DB 9; Length 524;
Best Local Similarity 52.9%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 72;

QY 8 TCAACCAAGACAGTCGCGATATCAAAATAAACAAGCAGCAATAGAGCCGGTGTGAACAA 67
DB 153 TTACAGGACCTGAACTACATGAACAACAACTGATGAACCTGAAAGAGAAATAGACAA 94
QY 68 CGCCGGATGCCAGACAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACCACCA 127
DB 93 TTCATAATTTAGTCAGAGACTTATATTCCTCTCTCCATAGTTAAACAGGGTAAAGTAGACA 34
QY 128 GAAATCGCCAGTGATATAAAGCAATTCAAAAA 160
DB 33 GAAATCAGTAACCTACAGAAACACATCAA 1

RESULT 14
AK085573/c

LOCUS
DEFINITION

AK085573 3190 bp mRNA linear HTC 03-APR-2004
Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D630042J08 product:unclassified, full insert
sequence.

ACCESSION
VERSION
AK085573.1 GI:26102821

KEYWORDS
HTC; CAP trapper.

SOURCE
Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
10349636

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

QY 153 TTACAGGACCTGAACTACATGAACAACAACTGATGAACCTGAAAGAGAAATAGACAA 94
QY 68 CGCCGGATGCCAGACAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACCACCA 127
DB 93 TTCATAATTTAGTCAGAGACTTATATTCCTCTCTCCATAGTTAAACAGGGTAAAGTAGACA 34
QY 128 GAAATCGCCAGTGATATAAAGCAATTCAAAAA 160
DB 33 GAAATCAGTAACCTACAGAAACACATCAA 1

AK085573 3190 bp mRNA linear HTC 03-APR-2004

Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched

library, clone:D630042J08 product:unclassified, full insert

sequence.

AK085573

AK085573.1 GI:26102821

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

9279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3190)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tgawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission

JOURNAL
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/

FEATURES

source
1. .3190
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D630042J08"
/db_xref="taxon:10090"
/clone="D630042J08"
/issue_type="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
misc_feature
1. .3190
/notes="unclassifiable"

ORIGIN

Query Match 9.6%; Score 37.8; DB 3; Length 3190;
Best Local Similarity 47.8%; Pred. No. 5.4;
Matches 111; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 148 GCAATTCAAAATAATCGGTGGGAAGCAGACGACGCCAAACACCATGTAAGTGAATGTTTC 1650
DB 1709 GGACTGAAAACATAGCATGCTGTATATACACCAACATGAATGTAAGTTC 1650

QY 208 TCATGTGACCTGGTTGTAAGAGAGAGACGCGGTATATTTTCAGCATCTGCGCG 267
DB 1649 TGAAGAGGACATTCGTAGACACAGAAAGTAAACCTGTATATTTCTCGTGTGG 1590

QY 268 CACAGAGGCGCATGGAAGCGCGGAGAGCAACATTCGTGTAGATTGATATTTAATATA 327
DB 1589 GAGTAAGGACATGGAGTACACACAGTTTACATAGTATGTAGNATATCTAGTCC 1530

QY 328 TTAGCGTAATCTGTATGCTGTTATCTATATATATGATCTAAATCACTTTAA 380
DB 1529 TCACATTATTTGTATGTATGTATGTATGACGAATATGAATGACTATAGTTTAA 1477

RESULT 15

AZ230411/c
LOCUS
DEFINITION
RPCI-23-69J14.TJ RPCI-23 Mus musculus genomic clone RPCI-23-69J14,
genomic survey sequence.

ACCESSION
AZ230411
VERSION
AZ230411.1 GI:8538457
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

TITLE
Mouse BAC End Sequences from Library RPCI-23

JOURNAL
Unpublished (1999)

COMMENT
Other GSSs: RPCI-23-69J14.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Reseach Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plates: 69 rows; J column: 14
Seq primer: SP6
Class: BAC ends.

FEATURES

source
1. .603
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-69J14"
/sex="female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site: 1; EcoRI; Site: 2; EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 9.5%; Score 37.2; DB 8; Length 603;
Best Local Similarity 53.4%; Pred. No. 5.8;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 210 ATGTTGACCTTGGTTGTAAAGAGAGAGAGCGCGCTTATTTTCAGCATCTGTCGCCA 269
DB 273 ATGTTGAAATGGATTTGTAGTCATATTGCAAAATGTTTTTGCCTGCAGCATCTAGTGCATC 214

QY 270 GAGAAGGCGATGAAAGCGCGGCGAGAGCAACATTGCTGTAGATTGATTTAATATATT 329
DB 213 AAAGTGCGAGGAGAAAGGAGGAGCAGTGCAGAAATTAACAAATCTCTGTAATAATA 154

QY 330 AGCGTAATCTTATGCTGTTATCTAT 355
DB 153 TGTGGAAGGTTAATGAATTAAGTTAT 128

Search completed: December 19, 2004, 02:47:03
Job time : 1716.72 secs

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 603)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.